

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 20:33:16 ; Search time 977.534 Seconds  
(without alignments)  
13479.231 Million cell updates/sec

Title: US-10-776-889-1

Perfect score: 2012

Sequence: 1 atcttagcccttgattata.....aactaagaagcaatacttca 2012

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2012	100.0	2012	19	US-10-776-889-1
2	2012	100.0	2012	19	US-10-776-311-3
3	2012	100.0	2012	22	US-10-985-109-81
4	2012	100.0	2012	22	US-10-985-254-81
c 5	1961.4	97.5	12456	22	US-10-985-109-95
c 6	1961.4	97.5	12456	22	US-10-985-254-95
7	1883	93.6	1883	19	US-10-776-889-13

RESULT 1  
US-10-776-889-1  
; Sequence 1, Application US/10776889  
; Publication No. US20040158052A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Zhan-Bin  
; APPLICANT: Kinney, Anthony  
; TITLE OF INVENTION: Seed Specific Promoters  
; FILE REFERENCE: BB1531 US NA  
; CURRENT APPLICATION NUMBER: US/10776.889  
; CURRENT FILING DATE: 2004-02-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 1  
; LENGTH: 2012  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-776-889-1

Query Match 100.0%; Score 2012; DB 19; Length 2012;  
Best Local Similarity 100.0%; Pred. No. 2e-242;  
Matches 2012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATCTTAGCCCTTGATTATATGTTGTAGATGCATGCAGTTTATTTCAAT	60
Db	1	ATCTTAGCCCTTGATTATATGTTGTAGATGCATGCAGTTTATTTCAAT	60
Qy	61	CCCTTTTCCCTTGAATCACTGACCACCAACCAAGAAAAAAGAAAGGATCA	120
Db	61	CCCTTTTCCCTTGAATCACTGACCACCAACCAAGAAAAAAGAAAGGATCA	120

Sequence 14, Appl  
Sequence 15, Appl  
Sequence 16, Appl  
Sequence 17, Appl  
Sequence 18, Appl  
Sequence 19, Appl  
Sequence 20, Appl  
Sequence 21, Appl  
Sequence 22, Appl  
Sequence 1, Appli  
Sequence 2128, Ap  
Sequence 386, App  
Sequence 386, App  
Sequence 986, App  
Sequence 240, App  
Sequence 255, App  
Sequence 255, App  
Sequence 769, App  
Sequence 240, App  
GENERAL INFORMATI  
Sequence 252, App  
Sequence 2369, Ap  
Sequence 230, App  
Sequence 1280, Ap  
Sequence 240, App  
Sequence 2331, Ap  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 113786,  
Sequence 113787,  
Sequence 113788,  
Sequence 113786,  
Sequence 113787,  
Sequence 113788,  
Sequence 239, App  
Sequence 33, Appl  
Sequence 6854, Ap

121 TTTTGAAGGATATTTTTGCGTCTCTATTTCAAACTACTGTATTTTTTACCAAAAACCTGTAT 180  
121 TTTTGAAGGATATTTTTGCGTCTCTATTTCAAACTACTGTATTTTTTACCAAAAACCTGTAT 180  
181 TTTTCTACACTCTCAAGCTTTGTTTTGCGTCTCGACTCTCATGATTTCTTTCATATGCC 240  
181 TTTTCTACACTCTCAAGCTTTGTTTTGCGTCTCGACTCTCATGATTTCTTTCATATGCC 240  
241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTTGCAAAAGCTTGTCAATCAA 300  
241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTTGCAAAAGCTTGTCAATCAA 300  
301 AGCTTGAATGTACAAATTAATGTTTTTCAATGCTTTCAAAATTTATCTGCACCCCTTAGC 360  
301 AGCTTGAATGTACAAATTAATGTTTTTCAATGCTTTCAAAATTTATCTGCACCCCTTAGC 360  
361 TATTAAATCTAACTCTAAAGTAAAGCTAGTGAATTTTTTTCGAATAGTCAATGCGTCAATTA 420  
361 TATTAAATCTAACTCTAAAGTAAAGCTAGTGAATTTTTTTCGAATAGTCAATGCGTCAATTA 420  
421 ATTTCCTCGTGAATTTTGGCTTTGACTCAACACACTGGCCCGGTACATCCGTCCTCAT 480  
421 ATTTCCTCGTGAATTTTGGCTTTGACTCAACACACTGGCCCGGTACATCCGTCCTCAT 480  
481 TACATGAAAAGAAATATTTGTTTATATTTCTTAATTTTAAATAATTTGTCCTTTCTAAATTTTC 540  
481 TACATGAAAAGAAATATTTGTTTATATTTCTTAATTTTAAATAATTTGTCCTTTCTAAATTTTC 540  
541 ATATAGTAAATTAATTAATTAATTTTCTTAATTTCTTAATTTCTAATTTTCAAAATTAAT 600  
541 ATATAGTAAATTAATTAATTTTCTTAATTTCTTAATTTCTAATTTTCAAAATTAAT 600  
601 ATTATGCTATGTAAAGTACATTAATTTTGGCTATATATCTTAAATTAATTTCTAAATTAAT 660  
601 ATTATGCTATGTAAAGTACATTAATTTTGGCTATATATCTTAAATTAATTTCTAAATTAAT 660  
661 TAAAAAAGCTGATATGAAAAATTTATTTCTTTTAAAGCTATATCATTTTTATATATAT 720  
661 TAAAAAAGCTGATATGAAAAATTTATTTCTTTTAAAGCTATATCATTTTTATATATAT 720  
721 TTTTCTTTCTTTCTTTCTAATTTCTAATTTCAATTTAAAGAAATAAATTTTGTAAATTT 780  
721 TTTTCTTTCTTTCTTTCTAATTTCTAATTTCAATTTAAAGAAATAAATTTTGTAAATTT 780  
781 TTATTTATCAATTTATAAATAATTTTACTTTATATTTTCTTTTCAATTTTGTGTAAACA 840  
781 TTATTTATCAATTTATAAATAATTTTACTTTATATTTTCTTTTCAATTTTGTGTAAACA 840  
841 AATCATATCATTTATGATTTGAAAGAGAGAAATTTGACAGTGTAGTAATTAAGTGTAGAGAAA 900  
841 AATCATATCATTTATGATTTGAAAGAGAGAAATTTGACAGTGTAGTAATTAAGTGTAGAGAAA 900  
901 AAATGTGTATTTCTTAAAAAAACCTTAAACAAACATGTATCTCTTATTTTCACTTA 960  
901 AAATGTGTATTTCTTAAAAAAACCTTAAACAAACATGTATCTCTTATTTTCACTTA 960  
961 TCTCTCAATTTCAATTTTCTTTATCTCTTTCTTTTATTTTTTATCATATCATTTTCACT 1020  
961 TCTCTCAATTTCAATTTTCTTTATCTCTTTCTTTTATTTTTTATCATATCATTTTCACT 1020  
1021 TAATTTATTTTACTCTCTTTATTTTCTCTATCTCTCTCTTTTATTTTCACTCATATAT 1080  
1021 TAATTTATTTTACTCTCTTTATTTTCTCTATCTCTCTCTTTTATTTTCACTCATATAT 1080  
1081 ACATCCAAAATTTGGGCGATGCGTTTATCATCTATCTCTCTCTCTCTCTCTCTCTCTCTCT 1140  
1081 ACATCCAAAATTTGGGCGATGCGTTTATCATCTATCTCTCTCTCTCTCTCTCTCTCTCTCT 1140  
1141 ATGAACTGAAAGCATTTGGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200  
1141 ATGAACTGAAAGCATTTGGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200

1201 TGGCATCTGATTTGATTTCAGTATATCTATTGTCATGTGTAAAGTCTTTTCCAAATACATAA 1260  
1201 TGGCATCTGATTTGATTTCAGTATATCTATTGTCATGTGTAAAGTCTTTTCCAAATACATAA 1260  
1261 CTATTAAATTAATCTTTAAATAAATAAAGGATAAATAATTTTTTTTTCTTCAAAAAATAAA 1320  
1261 CTATTAAATTAATCTTTAAATAAATAAAGGATAAATAATTTTTTTTTCTTCAAAAAATAAA 1320  
1321 ATATGTTATTTTTTTGTTTGTAGTGTATATTCGATATAATCTTAATATATATGATATGATTTT 1380  
1321 ATATGTTATTTTTTTGTTTGTAGTGTATATTCGATATAATCTTAATATATGATATGATTTT 1380  
1381 TTATATTTGATTTAAACATATAATCAATTAATTTAAATATGATATTTTTTATATAGTGTGACA 1440  
1381 TTATATTTGATTTAAACATATAATCAATTAATTTAAATATGATATTTTTTATATAGTGTGACA 1440  
1441 CATAAATTTTATAGGATAAATAATATGATAAATAAATAATTTTTTAAATATTTTTTATATTTAC 1500  
1441 CATAAATTTTATAGGATAAATAATATGATAAATAAATAATTTTTTAAATATTTTTTATATTTAC 1500  
1501 GAGAAAAAATAATTTTTAGCCATAAATAAATGACCCAGCATATTTTACACCTTAGTAAT 1560  
1501 GAGAAAAAATAATTTTTAGCCATAAATAAATGACCCAGCATATTTTACACCTTAGTAAT 1560  
1561 TCATAAAATTCCTATATGATATTTTGAATTTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620  
1561 TCATAAAATTCCTATATGATATTTTGAATTTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620  
1621 CTACGTCTCTCTGTCATTTGTTTTCATGCAAAAGAGGAGCGAAAAACCCCTCA 1680  
1621 CTACGTCTCTCTGTCATTTGTTTTCATGCAAAAGAGGAGCGAAAAACCCCTCA 1680  
1681 CCATGAATCACTCTTCCACACCATTTTTTACTAGCAAAACAGTCTCAACAACCTGAAGCCAGC 1740  
1681 CCATGAATCACTCTTCCACACCATTTTTTACTAGCAAAACAGTCTCAACAACCTGAAGCCAGC 1740  
1741 TCTCTTTCCGTTCTTTTTCACACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800  
1741 TCTCTTTCCGTTCTTTTTCACACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800  
1801 ATTTATTAACGTGTCGCAAAAGATGCTTTATTTGAATAGAGTGACATTTTGTATGTACTACTA 1860  
1801 ATTTATTAACGTGTCGCAAAAGATGCTTTATTTGAATAGAGTGACATTTTGTATGTACTACTA 1860  
1861 ATTTAGAACATGAAAAAGCATTTGTTCTAACACGATAATCTCTGTAAGGCGTTTAACTCCAAA 1920  
1861 ATTTAGAACATGAAAAAGCATTTGTTCTAACACGATAATCTCTGTAAGGCGTTTAACTCCAAA 1920  
1921 GATCCAAATTTCACTATATATAAATTTGTCAGAAAGCAAAATGAATTCATAGCTGAGAGAG 1980  
1921 GATCCAAATTTCACTATATATAAATTTGTCAGAAAGCAAAATGAATTCATAGCTGAGAGAG 1980  
1981 AAAGGAAAGGTTAACTTAAGCAATTAATTTCA 2012  
1981 AAAGGAAAGGTTAACTTAAGCAATTAATTTCA 2012

## RESULT 2

US-10-776-311-3

; Sequence 3, Application US/10776311

; Publication No. US20040172682A1

; GENERAL INFORMATION:

; APPLICANT: Edgar B. Cahoon

; APPLICANT: Howard G. Damude

; APPLICANT: William D. Hitz

; APPLICANT: Anthony J. Kinney

; APPLICANT: Charles W. Kolar

; APPLICANT: Zhan Bin Liu

; TITLE OF INVENTION: Production of Long Chain Polyunsaturated Fatty Acids in Plants

; FILE REFERENCE: BB1538 US NA

; CURRENT APPLICATION NUMBER: US/10/776,311

; CURRENT FILING DATE: 2004-02-11

; PRIOR APPLICATION NUMBER: US 60/446,941

```

; PRIOR FILING DATE: 2003-02-12
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1012
; TYPE: DNA
; ORGANISM: Glycine max
US-10-776-311-3

```

[illegible]

Qy	901	AAATGTTATTTCTCTAATAAATAACCTCAACAACATGATCTCTACTCTCTATTTTCATCTTA	960
Db	901	AAATGTTATTTCTCTAATAAATAACCTCAACAACATGATCTCTACTCTCTATTTTCATCTTA	960
Qy	961	TCCTCATTTTCATTTTCTCTTATCTCTTCTCTTCTTATTTTATCATATCATTTTCACAT	1020
Db	961	TCCTCATTTTCATTTTCTCTTATCTCTTCTTATTTTATCATATCATTTTCACAT	1020
Qy	1021	TAATATTTTTCATCTCTTTATTTTTTCTCTCTATCTCTATCCCTCTCTTATTTCCACTCATATAT	1080
Db	1021	TAATATTTTTCATCTCTTTATTTTTTCTCTCTATCTCTATCCCTCTCTTATTTCCACTCATATAT	1080
Qy	1081	ACACTCCAAAATTTGGGGCATGCGCTTTATCACTACTCTATCTCTCCCTCCACTCAATTCATTTAA	1140
Db	1081	ACACTCCAAAATTTGGGGCATGCGCTTTATCACTACTCTATCTCTCCCTCCACTCAATTCATTTAA	1140
Qy	1141	ATGAACCTGAAAAGCATTTGGCAAGTCTCTCCCTCTCTCAAGTGATTTCCAACTCAGCAT	1200
Db	1141	ATGAACCTGAAAAGCATTTGGCAAGTCTCTCCCTCTCTCAAGTGATTTCCAACTCAGCAT	1200
Qy	1201	TGGCATCTGATTCATTCAGTATATCTATTCGATGCTGTAAAAAGCTTTTCCACAATACATAA	1260
Db	1201	TGGCATCTGATTCATTCAGTATATCTATTCGATGCTGTAAAAAGCTTTTCCACAATACATAA	1260
Qy	1261	CTATTAATTAATCTTAAATAAATAAAGGATAAATAATTTTTTTCTTCATAAAAATTTAA	1320
Db	1261	CTATTAATTAATCTTAAATAAATAAAGGATAAATAATTTTTTTCTTCATAAAAATTTAA	1320
Qy	1321	ATATGTTATTTTTTGTGTAGATGATATATTCGAATAAATCTAAATATATGATATATGATTTT	1380
Db	1321	ATATGTTATTTTTTGTGTAGATGATATATTCGAATAAATCTAAATATATGATATATGATTTT	1380
Qy	1381	TTATATTTGATTAACATATAATCAATATTAATAATATGATATTTTTTTATATAGTTCGTACA	1440
Db	1381	TTATATTTGATTAACATATAATCAATATTAATAATATGATATTTTTTTATATAGTTCGTACA	1440
Qy	1441	CATAATTTTATAGGATAAAAAATATGATAAAAAATAAATTTTTTAAATATTTTATATTTTAC	1500
Db	1441	CATAATTTTATAGGATAAAAAATATGATAAAAAATAAATTTTTTAAATATTTTATATTTTAC	1500
Qy	1501	GAGAAAAAAAATATTTTAGCCATAAATAAATGACACAGCATATTTTTCAACACCTTAGTAAT	1560
Db	1501	GAGAAAAAAAATATTTTAGCCATAAATAAATGACACAGCATATTTTTCAACACCTTAGTAAT	1560
Qy	1561	TCATAAATTCCTATATGATATTTGAAATTTAAAAACAGATAATCGTTAAGGGAAGGAATC	1620
Db	1561	TCATAAATTCCTATATGATATTTGAAATTTAAAAACAGATAATCGTTAAGGGAAGGAATC	1620
Qy	1621	CTACGTCACTCTCTGCAATTTGTTTTTCATGCAAAACAGAAAGGACGAAACCACTCA	1680
Db	1621	CTACGTCACTCTCTGCAATTTGTTTTTCATGCAAAACAGAAAGGACGAAACCACTCA	1680
Qy	1681	CCATGAATCACTCTTACACCAATTTTTTACTAGCAACCAAGTCTCAACAACCTGAGCCAGC	1740
Db	1681	CCATGAATCACTCTTACACCAATTTTTTACTAGCAACCAAGTCTCAACAACCTGAGCCAGC	1740
Qy	1741	TCCTCTTCCGTTTTTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG	1800
Db	1741	TCCTCTTCCGTTTTTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG	1800
Qy	1801	ATTTATAACGTCGCAAAAGATGCTTATTTGAATAGAGTGCACATTTGTATATGTACTACTA	1860
Db	1801	ATTTATAACGTCGCAAAAGATGCTTATTTGAATAGAGTGCACATTTGTATATGTACTACTA	1860
Qy	1861	ATTTAGAACATGAAAAAGCATTTGTTCTTAAACAGTATATCTGTGAGGCGTTAACTCCAAA	1920
Db	1861	ATTTAGAACATGAAAAAGCATTTGTTCTTAAACAGTATATCTGTGAGGCGTTAACTCCAAA	1920
Qy	1921	GATCCAAATTTCACTATATAAATTTGTGCAAGAAACAAAATGAAATTCATAGCTGAGAGAG	1980
Db	1921	GATCCAAATTTCACTATATAAATTTGTGCAAGAAACAAAATGAAATTCATAGCTGAGAGAG	1980
Qy	1981	AAAGGAAAGGTTTAACTAAGAAAGCAATTACTTCA	2012

```
|||||
1981 AAAGAAAGGTTAACTAAGAACCAACTTCA 2012

RESULT 3
US-10-985-109-81
; Sequence 81, Application US/10985109
; Publication No. US20050132441A1
; GENERAL INFORMATION:
; APPLICANT: Yadav, Narendra
; APPLICANT: Damude, Howard
; TITLE OF INVENTION: DELTA-15 DESATURASES SUITABLE FOR ALTERING LEVELS OF POLYUNSATURATED FATTY ACIDS IN OILSEED PLANTS AND OLEAGINOUS YEAST
; FILE REFERENCE: CL2432
; CURRENT APPLICATION NUMBER: US/10/985.109
; CURRENT FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/519191
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81
; LENGTH: 2012
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-985-109-81

Query Match 100.0%; Score 2012; DB 22; Length 2012;
Best Local Similarity 100.0%; Pred. No. 2e-242; Mismatches 0; Indels 0; Gaps 0;
Matches 2012; Conservative 0;

Qy 1 ATCTTAGGCCCTTGATTATATGTTGGTTTAGATGATTACATGCAAGTTTTATTTCAAT 60
Db 1 ATCTTAGGCCCTTGATTATATGTTGGTTTAGATGATTACATGCAAGTTTTATTTCAAT 60

Qy 61 CCCTTTTCCTTTGAATACGACCAAGAACCAAGAAAAAAGAAAAAGGATCA 120
Db 61 CCCTTTTCCTTTGAATACGACCAAGAACCAAGAAAAAAGAAAAAGGATCA 120

Qy 121 TTTTGAAGGATATTTTCGCTCTATTCAATACTGTATTTTACCAAAAAAACTGTAT 180
Db 121 TTTTGAAGGATATTTTCGCTCTATTCAATACTGTATTTTACCAAAAAAACTGTAT 180

Qy 181 TTTTCTACACTCTCAAGCTTTGTTTTGCGTTCGACTCTCATGATTTCTTCATATGCC 240
Db 181 TTTTCTACACTCTCAAGCTTTGTTTTGCGTTCGACTCTCATGATTTCTTCATATGCC 240

Qy 241 AATCACTCTATTTATAAATGSCATAAGTAGTGGAACAATTCGAAGCTTGTCAACAAA 300
Db 241 AATCACTCTATTTATAAATGSCATAAGTAGTGGAACAATTCGAAGCTTGTCAACAAA 300

Qy 301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCTTTTCAAAATTCATGCAACCCCTAGC 360
Db 301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCTTTTCAAAATTCATGCAACCCCTAGC 360

Qy 361 TATTAATCTAACACTCAAGTAGGCTAGTGAATTTTTTCGAATAGTCATGCGATGCAATTA 420
Db 361 TATTAATCTAACACTCAAGTAGGCTAGTGAATTTTTTCGAATAGTCATGCGATGCAATTA 420

Qy 421 ATTTCCCGTGACTATTTTGGCTTTGACTCCAAACACTGGCCCGGTACATCGGTCCTCAT 480
Db 421 ATTTCCCGTGACTATTTTGGCTTTGACTCCAAACACTGGCCCGGTACATCGGTCCTCAT 480

Qy 481 TACATGAAGAAATATTTGTTTATTTATTTTAAATAATTCGTCCTTCTAAATTTTC 540
Db 481 TACATGAAGAAATATTTGTTTATTTATTTTAAATAATTCGTCCTTCTAAATTTTC 540

Qy 541 ATATAGTTAAATTAATTAATTAATTTTCTCTATTTCTATTTAGTTCTATTTTCAAAATTAAT 600
Db 541 ATATAGTTAAATTAATTAATTAATTTTCTCTATTTCTATTTAGTTCTATTTTCAAAATTAAT 600

Qy 601 ATTTATGCATATGTAAGTACATTTATTTTGGCTATATTTTGTCTATATTTTCTAAATTAAT 660
Db 601 ATTTATGCATATGTAAGTACATTTATTTTGGCTATATTTTGTCTATATTTTCTAAATTAAT 660
```

```
Qy 661 TAAAAAAGACTCATATGAAAAAATTTATCTTTTAAAGCTATATCATTTTATATATACT 720
Db 661 TAAAAAAGACTCATATGAAAAAATTTATCTTTTAAAGCTATATCATTTTATATATACT 720

Qy 721 TTTTCTTTTCTTTTCTTTTCAATTTTCAATTTAATTAAGAAATAAATTTTGTAAATTT 780
Db 721 TTTTCTTTTCTTTTCTTTTCAATTTTCAATTTAATTAAGAAATAAATTTTGTAAATTT 780

Qy 781 TTATTTTCAATTTTATAAAAATTTTACTTTTATATGTTTTTTCACATTTTGTGTAACA 840
Db 781 TTATTTTCAATTTTATAAAAATTTTACTTTTATATGTTTTTTCACATTTTGTGTAACA 840

Qy 841 AATCATATCATTTATGATTTGAAAGAGAGAAATTTGACAGTGAGTAATAAGTGATGAGAAA 900
Db 841 AATCATATCATTTATGATTTGAAAGAGAGAAATTTGACAGTGAGTAATAAGTGATGAGAAA 900

Qy 901 AAATGTTTATTTTCTTAAAAAACCTTAAACAAACATGTATCTCTTATTTTCATCTA 960
Db 901 AAATGTTTATTTTCTTAAAAAACCTTAAACAAACATGTATCTCTTATTTTCATCTA 960

Qy 961 TCTCTCATTTTCTTTTCTTTTATCTCTTTCTTTTATTTTATTTTATCATATCATTTCCAT 1020
Db 961 TCTCTCATTTTCTTTTCTTTTATCTCTTTTATCTCTTTTATTTTATCATATCATTTCCAT 1020

Qy 1021 TAATTTATTTTACTCTCTTTTATTTTCTCTCTATCCCTCTCTTATTTTCCACTCATATAT 1080
Db 1021 TAATTTATTTTACTCTCTTTTATTTTCTCTCTATCCCTCTCTTATTTTCCACTCATATAT 1080

Qy 1081 ACACCTCAAAATTTGGGGCATGCTTATCACTACTCTATCTCTCCCTCAAAATCAATTTAA 1140
Db 1081 ACACCTCAAAATTTGGGGCATGCTTATCACTACTCTATCTCTCCCTCAAAATCAATTTAA 1140

Qy 1141 ATGAAACTGAAAGACATTTGGCAAGTCTCTCTCCCTCTCAAGTGATTTTCCACTCAGCAT 1200
Db 1141 ATGAAACTGAAAGACATTTGGCAAGTCTCTCTCCCTCTCAAGTGATTTTCCACTCAGCAT 1200

Qy 1201 TGGCATCTGATTTGATTTGAGTATATCTTATGTCATGTGTAAGTCTTTCCACAATACATAA 1260
Db 1201 TGGCATCTGATTTGATTTGAGTATATCTTATGTCATGTGTAAGTCTTTCCACAATACATAA 1260

Qy 1261 CTATTAATTAATCTTAAAAATAAGGATAAATAATTTTCTTCTCAATAAAATTTAA 1320
Db 1261 CTATTAATTAATCTTAAAAATAAGGATAAATAATTTTCTTCTCAATAAAATTTAA 1320

Qy 1321 ATATGTTATTTTGTGTTAGATGATATTCGAATAATCTTAAATATATGATATGATTTT 1380
Db 1321 ATATGTTATTTTGTGTTAGATGATATTCGAATAATCTTAAATATATGATATGATTTT 1380

Qy 1381 TTATATTTGATTTAAACATATAATCAATATTTAAATATGATATTTTATATAGTTGTACA 1440
Db 1381 TTATATTTGATTTAAACATATAATCAATATTTAAATATGATATTTTATATAGTTGTACA 1440

Qy 1441 CATATAATTTTAAAGGATAAATAATGATAAATAAATTTTAAATATTTTATATTTTAC 1500
Db 1441 CATATAATTTTAAAGGATAAATAATGATAAATAAATTTTAAATATTTTATATTTTAC 1500

Qy 1501 GAGAAAAAATAATTTTAGCCATAAATAAATGACCCAGCATATTTTACAACTTTAGTAAT 1560
Db 1501 GAGAAAAAATAATTTTAGCCATAAATAAATGACCCAGCATATTTTACAACTTTAGTAAT 1560

Qy 1561 TCATAAAATCTCTATATGATATTTGAAATTTAAAAACAGATAATTCGTTTAAAGGGAAGATC 1620
Db 1561 TCATAAAATCTCTATATGATATTTGAAATTTAAAAACAGATAATTCGTTTAAAGGGAAGATC 1620

Qy 1621 CTACGTCTCTCTTGGCCATTTGTTTTTCATGAAAACAGAAAGGAGCAAAACCACTCA 1680
Db 1621 CTACGTCTCTCTTGGCCATTTGTTTTTCATGAAAACAGAAAGGAGCAAAACCACTCA 1680

Qy 1681 CCATGAATCACTCTTCCACACCATTTTCTAGCAAAACAGTCTCAACAACTGAAGCCAGC 1740
Db 1681 CCATGAATCACTCTTCCACACCATTTTCTAGCAAAACAGTCTCAACAACTGAAGCCAGC 1740
```

```
QY 1741 TCTCTTCGGTTCCTTTTACAACTTCTTGAATAGTAGTATTTTTTTTTCACATG 1800
Db 1741 TCTCTTCGGTTCCTTTTACAACTTCTTGAATAGTAGTATTTTTTTTTCACATG 1800
QY 1801 ATTTATTAAACGTCGCAAAAGATGCTTAATGAATAGAGTGCAATTTGTAATGTACTACTA 1860
Db 1801 ATTTATTAAACGTCGCAAAAGATGCTTAATGAATAGAGTGCAATTTGTAATGTACTACTA 1860
QY 1861 ATTGAACATGAAAAAGCATTGTTCTTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920
Db 1861 ATTGAACATGAAAAAGCATTGTTCTTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920
QY 1921 GATCCAATTCACATATATAATTTGACGAAGCAAAATGAATTCACATAGCTGAGAG 1980
Db 1921 GATCCAATTCACATATATAATTTGACGAAGCAAAATGAATTCACATAGCTGAGAG 1980
QY 1981 AAAGGAAAGGTTAACTAAGAACCAATCTTCA 2012
Db 1981 AAAGGAAAGGTTAACTAAGAACCAATCTTCA 2012

RESULT 4
US-10-985-254-81
; Sequence 81, Application US/10985254
; Publication No. US20050132442A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Co., Inc.
; APPLICANT: Yadav, Narendra
; APPLICANT: Damude, Howard
; TITLE OF INVENTION: DELTA 15 DESATURASES SUITABLE FOR ALTERING LEVELS OF
; TITLE OF INVENTION: POLYUNSATURATED FATTY ACIDS IN OLEAGINOUS YEAST
; FILE REFERENCE: CL2432
; CURRENT APPLICATION NUMBER: US/10/985,254
; CURRENT FILING DATE: 2004-11-10
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patencin version 3.3
; SEQ ID NO 81
; LENGTH: 2012
; TYPE: DNA
; ORGANISM: Glycine max
US-10-985-254-81

Query Match 100.0%; Score 2012; DB 22; Length 2012;
Best Local Similarity 100.0%; Pred. No. 2e-242;
Matches 2012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTAGGCCCTTGATTTATATATGGTGTGTAGATGGATTCACATGCAGTTTTTATTTCAAT 60
Db 1 ATCTTAGGCCCTTGATTTATATATGGTGTGTAGATGGATTCACATGCAGTTTTTATTTCAAT 60
QY 61 CCCTTTTCTTTGTAATACTGACCAAGAAACAAGAAAAAAGAAAAAGGATCA 120
Db 61 CCCTTTTCTTTGTAATACTGACCAAGAAACAAGAAAAAAGAAAAAGGATCA 120
QY 121 TTTTGAAGGATATTTTTCGCTCCTATTCAAAATCTGTATTTTACCAAAAAAACTGTAT 180
Db 121 TTTTGAAGGATATTTTTCGCTCCTATTCAAAATCTGTATTTTACCAAAAAAACTGTAT 180
QY 181 TTTTCTACACTCTCAAGCTTTGTTTTCGCTTGGACTCTCATGTTTCTTCATATGCC 240
Db 181 TTTTCTACACTCTCAAGCTTTGTTTTCGCTTGGACTCTCATGTTTCTTCATATGCC 240
QY 241 AATCACTCTATTATATAATGCGATAGGTAGTGGAACAATTTGCAAGCTTTGTGCATCAA 300
Db 241 AATCACTCTATTATATAATGCGATAGGTAGTGGAACAATTTGCAAGCTTTGTGCATCAA 300
QY 301 AGCTTGAAGTACAAATTAATGTTTTTCATGCCCTTTTCAAAATTAATCTGCACCCCTAGC 360
Db 301 AGCTTGAAGTACAAATTAATGTTTTTCATGCCCTTTTCAAAATTAATCTGCACCCCTAGC 360
QY 361 TATTAATCTAATCATCTAAGTAAGGCTAGTGAATTTTTTTCGAATAGTCATGCGATGCAATTA 420
```

```
Db 361 TATTAATCTAATCATCTAAGTAAGGCTAGTGAATTTTTTTCGAATAGTCATGCGATGCAATTA 420
QY 421 ATTTCCCGCGTACATATTTTGGCTTTTGACTCCAAACACTGGCCCCCGTACATCCGTCCTCAT 480
Db 421 ATTTCCCGCGTACATATTTTGGCTTTTGACTCCAAACACTGGCCCCCGTACATCCGTCCTCAT 480
QY 481 TACATGAAAAAGAAATATTTGTTTATATTTCTTAATTTAAATAATTTGTCCTCTCTAAATTTTC 540
Db 481 TACATGAAAAAGAAATATTTGTTTATATTTCTTAATTTAAATAATTTGTCCTCTCTAAATTTTC 540
QY 541 ATATAGTTAAATATATATATTTTCTCTATTTCTATTTCTATTTAGTTCTATTTTCAAAATTTT 600
Db 541 ATATAGTTAAATATATATATTTTCTCTATTTCTATTTCTATTTAGTTCTATTTTCAAAATTTT 600
QY 601 ATTTATGCATATGTAAAGTACATTTATATTTTGTCTATATCTTAAATATTTTCTAAATTTAT 660
Db 601 ATTTATGCATATGTAAAGTACATTTATATTTTGTCTATATCTTAAATATTTTCTAAATTTAT 660
QY 661 TAAAAAAGACTGATATGAAAAATTTATTTCTTTTAAAGCTATATCATTTTATATATACT 720
Db 661 TAAAAAAGACTGATATGAAAAATTTATTTCTTTTAAAGCTATATCATTTTATATATACT 720
QY 721 TTTTCTTTTCTTTTCTTTTCTATTTCTCAATTTTAAAGAAATAAATTTTGTAAATTTT 780
Db 721 TTTTCTTTTCTTTTCTTTTCTATTTCTCAATTTTAAAGAAATAAATTTTGTAAATTTT 780
QY 781 TTATTTATCAATTTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTGTAAAAACA 840
Db 781 TTATTTATCAATTTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTGTAAAAACA 840
QY 841 AATCATATCATATGATTTGAAAGAGAGAAATTTGACAGTGAGTAATAGTGATGAGAAAA 900
Db 841 AATCATATCATATGATTTGAAAGAGAGAAATTTGACAGTGAGTAATAGTGATGAGAAAA 900
QY 901 AAATGCTGTATTTCTTAAAAAACCTTAAACAAACATGTATCTACTCTCTATTTTCATCTA 960
Db 901 AAATGCTGTATTTCTTAAAAAACCTTAAACAAACATGTATCTACTCTCTATTTTCATCTA 960
QY 961 TCTCTCAATTTCTCTTTTATCTCTTTTCTTTTATTTTATCATATCATTTTCACAT 1020
Db 961 TCTCTCAATTTCTCTTTTCTCTTTTATCTCTTTTATTTTATCATATCATTTTCACAT 1020
QY 1021 TAATATTTTATCTCTCTTTTATTTTCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTATAT 1080
Db 1021 TAATATTTTATCTCTCTTTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTATAT 1080
QY 1081 ACCTCCAAATTTGGGGCATGCCCTTTTATCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
Db 1081 ACCTCCAAATTTGGGGCATGCCCTTTTATCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
QY 1141 ATGAAACTGAAAAAGCATTTGGCAAGTCTCTCTCCCTCCCTCAAGTGATTTTCCAACTCAGCAT 1200
Db 1141 ATGAAACTGAAAAAGCATTTGGCAAGTCTCTCTCCCTCCCTCAAGTGATTTTCCAACTCAGCAT 1200
QY 1201 TGGCATCTGATTTGATTCAGTATATCTATTTGATGATGTGTAAGTCTTTCCCAATACATAA 1260
Db 1201 TGGCATCTGATTTGATTCAGTATATCTATTTGATGATGTGTAAGTCTTTCCCAATACATAA 1260
QY 1261 CTATTAATTAATCTTAAATAAAGGATAAATAATTTTTTTTCTTCTTCAAAAAATTTAA 1320
Db 1261 CTATTAATTAATCTTAAATAAAGGATAAATAATTTTTTTTCTTCTTCAAAAAATTTAA 1320
QY 1321 ATATGTTATTTTGTGTTAGATGTATATTCGAATAAATCTAAATATATATGATAATGATTTT 1380
Db 1321 ATATGTTATTTTGTGTTAGATGTATATTCGAATAAATCTAAATATATATGATAATGATTTT 1380
QY 1381 TTATTTGATTAACAATATATTAATCAATTAATAATGATATTTTTTTTATATAGTTGTACA 1440
Db 1381 TTATTTGATTAACAATATATTAATCAATTAATAATGATATTTTTTTTATATAGTTGTACA 1440
QY 1441 CATAAATTTTATAGGATAAATAATGATAAATAAATTTTAAATATTTTATATTTTATTTAC 1500
```

```
Db 1441 CATAAATTTTATAGGATAAAAAATATGATAAAAAATAAATTTTAAATATATTTTATATTTAC 1500
Qy 1501 GAGAAAAAATAATTTTAGCATAAAATAAATGACACAGCATATTTTCAACCTTAGTAAT 1560
Db 1501 GAGAAAAAATAATTTTAGCATAAATAAATGACACAGCATATTTTCAACCTTAGTAAT 1560
Qy 1561 TCATAAATTCCTATATATATATTTGAAATTAATAAACAGATAATCGTTAAGGGAAGGAATC 1620
Db 1561 TCATAAATTCCTATATATATATTTGAAATTAATAAACAGATAATCGTTAAGGGAAGGAATC 1620
Qy 1621 CTAGCTCATCTCTGCCATTTGTTTTTCATGCAAAACAGAAAGGACGAAAAACCACTCA 1680
Db 1621 CTAGCTCATCTCTGCCATTTGTTTTTCATGCAAAACAGAAAGGACGAAAAACCACTCA 1680
Qy 1681 CCATGAATCACTCTTCACACCATTTTCTAGCAAAACAGAGTCTCAACAACTGGAAGCCAGC 1740
Db 1681 CCATGAATCACTCTTCACACCATTTTCTAGCAAAACAGAGTCTCAACAACTGGAAGCCAGC 1740
Qy 1741 TCTCTTCCGTTCTTTTTACAACACTTCTTTTGAATAGTAGTATTTTTTTTTCACATG 1800
Db 1741 TCTCTTCCGTTCTTTTTACAACACTTCTTTTGAATAGTAGTATTTTTTTTTCACATG 1800
Qy 1801 ATTTTAAAGTGCACAAAGATGCTTATGATAGAGTGACATTTGCTAATGTAAGTACTA 1860
Db 1801 ATTTTAAAGTGCACAAAGATGCTTATGATAGAGTGACATTTGCTAATGTAAGTACTA 1860
Qy 1861 ATTAGAACATGAAAAAGCATTTGTTCTAACACGATAATCCTGTGAAGCGTTAACTCCAAA 1920
Db 1861 ATTAGAACATGAAAAAGCATTTGTTCTAACACGATAATCCTGTGAAGCGTTAACTCCAAA 1920
Qy 1921 GATCCAATTTCACTATATAAATTTGTGACGAAGCAAAATGAATTCACATAGCTGAGAGAG 1980
Db 1921 GATCCAATTTCACTATATAAATTTGTGACGAAGCAAAATGAATTCACATAGCTGAGAGAG 1980
Qy 1981 AAAGGAAGGTTAACTAAGCAAGCAATCTCA 2012
Db 1981 AAAGGAAGGTTAACTAAGCAAGCAATCTCA 2012

RESULT 5
US-10-985-109-95/c
; Sequence 95, Application US/10985109
; Publication No. US20050132441A1
; GENERAL INFORMATION:
; APPLICANT: Yadav, Narendra
; APPLICANT: Damude, Howard
; TITLE OF INVENTION: DELTA-15 DESATURASES SUITABLE FOR ALTERING LEVELS OF POLYUNSATURATED FATTY ACIDS IN OILSEED PLANTS AND OLEAGINOUS YEAST
; FILE REFERENCE: CL2432
; CURRENT APPLICATION NUMBER: US/10/985,109
; CURRENT FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/519191
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 95
; LENGTH: 12456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pK585
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1201)..(1201)
; OTHER INFORMATION: n is a, c, g, or t

US-10-985-109-95
Query Match 97.5%; Score 1961.4; DB 22; Length 12456;
Best Local Similarity 99.8%; Pred. No. 6.5e-236;
Matches 1974; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
Qy 38 CACATGCAAGTTTTTATTTCAATCCCTTTTCCCTTGAATTAACATGACCAAGAACACAG- 96
|||||
```

```
Db 12431 CACATGCAAGTTTTTATTTCAATCCCTTTTCTTTGAATAACTGACCAAGAACACAAGA 12372
Qy 97 -AAAAAAGGATCATTTTGAAGGATATTTTTCGCTCCTTATTTCAAATAC 155
Db 12371 AAAAAAAGGATCATTTTGAAGGATATTTTTCGCTCCTTATTTCAAATAC 12312
Qy 156 TGTATTTTACCAAAAAAAGTATTTTCTCACTCTCAAGCTTTGTTTTCGCTTCG 215
Db 12311 TGTATTTTACCAAAAAAAGTATTTTCTCACTCTCAAGCTTTGTTTTCGCTTCG 12252
Qy 216 ACTCTCATGATTTCTTATATGCGCAATCACTCTTATTAATGCGATAGTAGTG 275
Db 12251 ACTCTCATGATTTCTTATGCGCAATCACTCTTATTAATGCGATAGTAGTG 12192
Qy 276 AACAAATTGCAAGCTTGTGATCAAAAGCTTGCATGACAAATTAATGTTTTCATGCGCT 335
Db 12191 AACAAATTGCAAGCTTGTGATCAAAAGCTTGCATGACAAATTAATGTTTTCATGCGCT 12132
Qy 336 TTCAAAATTTATCTGCAACCCCTAGCTATTAATCTAACTCAAGTAAAGGCTAGTGAATTT 395
Db 12131 TTCAAAATTTATCTGCAACCCCTAGCTATTAATCTAACTCAAGTAAAGGCTAGTGAATTT 12072
Qy 396 TTTGGAATAGTCAAGTGCATTAATTTCCCGTGACTATTTTGGCTTTGACTCCAACA 455
Db 12071 TTTGGAATAGTCAAGTGCATTAATTTCCCGTGACTATTTTGGCTTTGACTCCAACA 12012
Qy 456 CTGGCCCCGTACATCCGTCCTCATTAATGACAAAGAAATATTTGTTTATATCTTAATTA 515
Db 12011 CTGGCCCCGTACATCCGTCCTCATTAATGACAAAGAAATATTTGTTTATATCTTAATTA 11952
Qy 516 AAAAATTTGTCCTTCTAAATTTTCATATAGTTAAATTTATATATTAATTTTCTCTATT 575
Db 11951 AAAAATTTGTCCTTCTAAATTTTCATATAGTTAAATTTATATATTAATTTTCTCTATT 11892
Qy 576 CTATAGTTCTATTTTCAAAATTTATATATGATGTAAGTACATATATTTTGTCT 635
Db 11891 CTATAGTTCTATTTTCAAAATTTATATATGATGTAAGTACATATATTTTGTCT 11832
Qy 636 ATATACCTAAATTTTCTAAATTTTAAAGGAGCTGATATGAAAAATTTATTTCTTTT 695
Db 11831 ATATACCTAAATTTTCTAAATTTTAAAGGAGCTGATATGAAAAATTTATTTCTTTT 11772
Qy 696 AAAGCTATATCAATTTTATATATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCAATTT 755
Db 11771 AAAGCTATATCAATTTTATATATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCAATTT 11712
Qy 756 AATAAGAAATAAATTTTGTAAAAATTTTATTTATCAATTTTATAAAAAATTTTACTTTATA 815
Db 11711 AATAAGAAATAAATTTTGTAAAAATTTTATTTATCAATTTTATAAAAAATTTTACTTTATA 11652
Qy 816 TGTTTTTTCCACATTTTGTAAACAAATCATATCATTTATGATTGAAAGAGAGAAATTTGA 875
Db 11651 TGTTTTTTCCACATTTTGTAAACAAATCATATCATTTATGATTGAAAGAGAGAAATTTGA 11592
Qy 876 CAGTGAGTAAATAGTGATGAGAAAAAATGTGTTATTTCTTAAAAAACCTTAAACAAAC 935
Db 11591 CAGTGAGTAAATAGTGATGAGAAAAAATGTGTTATTTCTTAAAAAACCTTAAACAAAC 11532
Qy 936 ATGTATCTACTCTCTATTTTCATCTATCTCTCATTTCTTTTCTTTTCTTTTCTTTTCTTTT 995
Db 11531 ATGTATCTACTCTCTATTTTCATCTATCTCTCATTTCTTTTCTTTTCTTTTCTTTTCTTTT 11472
Qy 996 ATTTTTTTATCATATCATTTTCAATTAATTTTATTTTACTCTCTCTTTTATTTTCTCTCTAT 1055
Db 11471 ATTTTTTTATCATATCATTTTCAATTAATTTTATTTTACTCTCTCTTTTATTTTCTCTCTAT 11412
Qy 1056 CCGCTCTCTTTTCCACTCATATATACACTCCAAATTTGGGCGATCGCTTTTATCACTACT 1115
Db 11411 CCGCTCTCTTTTCCACTCATATATACACTCCAAATTTGGGCGATCGCTTTTATCACTACT 11352
Qy 1116 CTATCTCTCTCTAAATCATTTTAATGAACTGAAAGCAATTTGGCAAGCTTCTCTCCCT 1175
Db 11351 CTATCTCTCTCTAAATCATTTTAATGAACTGAAAGCAATTTGGCAAGCTTCTCTCCCT 11292
```

```
QY 1176 CCTCAAGTGATTTCCAACTCAGCATTTGGCATCTGATTGATTCAGTATATCTATTCGATGT 1235
DB 11291 CCTCAAGTGATTTCCAACTCAGCATTTGGCATCTGATTGATTCAGTATATCTATTCGATGT 11232
QY 1236 GTAAAGTCTTTCCCAATACATACTATTAATAATCTTAATAATAATAAGGATAAAAT 1295
DB 11231 GTAAAGTCTTTCCCAATACATACTATTAATAATCTTAATAATAATAAGGATAAAAT 11172
QY 1296 ATTTTTTTTCTTCATAAAATTAATAATATGTTATTTTTTGTGTTAGATGATATTCGAATA 1355
DB 11171 ATTTTTTTTCTTCATAAAATTAATAATATGTTATTTTTTGTGTTAGATGATATTCGAATA 11112
QY 1356 AATCTAAATATATGATATGATTTTTTTATATGATTAACATATATCAATATTAATAT 1415
DB 11111 AATCTAAATATATGATATGATTTTTTTATATGATTAACATATATCAATATTAATAT 11052
QY 1416 GATATTTTTTTATATAGTTGTGACACATAATTTTATAAGGATAAAAAATATGATATAAAAT 1475
DB 11051 GATATTTTTTTATATAGTTGTGACACATAATTTTATAAGGATAAAAAATATGATATAAAAT 10992
QY 1476 AAATTTTAAATATTTTTTATATACGAGAAAAAAATATTTTAGCCATAAATAAATGAC 1535
DB 10991 AAATTTTAAATATTTTTTATATACGAGAAAAAAATATTTTAGCCATAAATAAATGAC 10932
QY 1536 CAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGATATATTTGAAATTTAAAAA 1595
DB 10931 CAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGATATATTTGAAATTTAAAAA 10872
QY 1596 CAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTGGCATTTGTTTTTCATGCAAA 1655
DB 10871 CAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTGGCATTTGTTTTTCATGCAAA 10812
QY 1656 CAGAAGGAGCGAAAAACCACTCACCATGAATCACTCTTCACACCAATTTTACTAGCAA 1715
DB 10811 CAGAAGGAGCGAAAAACCACTCACCATGAATCACTCTTCACACCAATTTTACTAGCAA 10752
QY 1716 ACAAGTCTCAACACTGAAGCAGCTCTCTTCGTTCTTTCTTTTACAACACTTTCTTTGA 1775
DB 10751 ACAAGTCTCAACACTGAAGCAGCTCTCTTCGTTCTTTCTTTTACAACACTTTCTTTGA 10692
QY 1776 AATAGTAGTATTTTTTTTTCATCATGATTTTAACTGCGCAAAAGATGCTTATTGAATAG 1835
DB 10691 AATAGTAGTATTTTTTTTTCATCATGATTTTAACTGCGCAAAAGATGCTTATTGAATAG 10632
QY 1836 AGTCACATTTGTATGTACTACTAATTAGAACATGAAAAAGCAATTTCTTAACACGATA 1895
DB 10631 AGTCACATTTGTATGTACTACTAATTAGAACATGAAAAAGCAATTTCTTAACACGATA 10572
QY 1896 ATCTGTGAAGCGTTAACTCCAAGATCCAAATTTCACTATATATAAATTTGTGACGAAAGCA 1955
DB 10571 ATCTGTGAAGCGTTAACTCCAAGATCCAAATTTCACTATATATAAATTTGTGACGAAAGCA 10512
QY 1956 AAATCAATTCACATAGCTGAGAGAGAAAGGAAGTTAACTAAGAACCAATCTCA 2012
DB 10511 AAATCAATTCACATAGCTGAGAGAGAAAGGAAGTTAACTAAGAACCAATCTCA 10455
```

## RESULT 6

```
US-10-985-254-95/c
; Sequence 95, Application US/10985254
; Publication No. US20050132442A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Co., Inc.
; APPLICANT: Yadav, Narendra
; APPLICANT: Damude, Howard
; TITLE OF INVENTION: DELTA 15 DESATURASES SUITABLE FOR ALTERING LEVELS OF
; FILE OF INVENTION: POLYUNSATURATED FATTY ACIDS IN OLEAGINOUS YEAST
; FILE REFERENCE: CL2432
; CURRENT APPLICATION NUMBER: US/10/985,254
; CURRENT FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/519191
; PRIOR FILING DATE: 2003-11-12
```

```
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 95
; LENGTH: 12456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pKRS85
; NAME/KEY: misc feature
; LOCATION: (1201)..(1201)
; OTHER INFORMATION: n is a, c, g, or t
US-10-985-254-95
```

```
Query Match 97.5%; Score 1961.4; DB 22; Length 12456;
Best Local Similarity 99.8%; Pred. No. 6.5e-236;
Matches 1974; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 38 CACATGCAAGTTTTTATTTCAATCCCTTTTCCCTTTTGAATACTGACCAAGAACACACAG- 96
DB 12431 CACATGCAAGTTTTTATTTCAATCCCTTTTCCCTTTTGAATACTGACCAAGAACACACAGA 12372
QY 97 -AAAAAAGAAAAAGAAAGGATCATTTTGAAGGATATTTTTCGCTCTCTATTCAAATAC 155
DB 12371 AAAAAAAGAAAAAGAAAGGATCATTTTGAAGGATATTTTTCGCTCTCTATTCAAATAC 12312
QY 156 TGTATTTTTTACCACAAAAAAGTATTTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCG 215
DB 12311 TGTATTTTTTACCACAAAAAAGTATTTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCG 12252
QY 216 ACTCTCATGATTTTCTTCATATGCAATCACTCTATTTATTAATGGCATAGGTAGTGTG 275
DB 12251 ACTCTCATGATTTTCTTCATATGCAATCACTCTATTTATTAATGGCATAGGTAGTGTG 12192
QY 276 AACAAATTTGCAAGCTTTGTTCATCAAAAGCTTGCATATGACAAATTAATGTTTTTCATGCT 335
DB 12191 AACAAATTTGCAAGCTTTGTTCATCAAAAGCTTGCATATGACAAATTAATGTTTTTCATGCT 12132
QY 336 TTCAAAATTTCTGCACCCCTAGCTATTAATTAACATCTAAGTAAGGCTAGTGAATTT 395
DB 12131 TTCAAAATTTCTGCACCCCTAGCTATTAATTAACATCTAAGTAAGGCTAGTGAATTT 12072
QY 396 TTTTCAATGATGATGATGATGATTAATTTCCCTGCTGATTTTGGCTTTGACCTCCACAC 455
DB 12071 TTTTCAATGATGATGATGATTAATTTCCCTGCTGATTTTGGCTTTGACCTCCACAC 12012
QY 456 CTGGCCCGGTACATCCCTGCTCATTAATGACAAAGAAATATGTTTATATATTTCTTAATTA 515
DB 12011 CTGGCCCGGTACATCCCTGCTCATTAATGACAAAGAAATATGTTTATATATTTCTTAATTA 11952
QY 516 AAAATATTTGCTCCCTTTCTAAATTTTTCATATAGTATTAATTTATATATTTCTCTATT 575
DB 11951 AAAATATTTGCTCCCTTTCTAAATTTTTCATATAGTATTAATTTATATATTTCTCTATT 11892
QY 576 CTATAGTCTCTATTTTCAAAATTTATTTATGATATGATAAGTACATATATTTTGTCT 635
DB 11891 CTATAGTCTCTATTTTCAAAATTTATTTATGATATGATAAGTACATATATTTTGTCT 11832
QY 636 ATATACCTTAAATATTTCTAAATTTTAAAGAAAGACATGATATGAAAAATTTATTTCTTTT 695
DB 11831 ATATACCTTAAATATTTCTAAATTTTAAAGAAAGACATGATATGAAAAATTTATTTCTTTT 11772
QY 696 AAAGCTATATCATTTTATATATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 755
DB 11771 AAAGCTATATCATTTTATATATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 11712
QY 756 AATGAAGAAATTAATTTTGTAAATTTTATTTATCAATTTTATTAATAATTTTACTTTATA 815
DB 11711 AATGAAGAAATTAATTTTGTAAATTTTATTTATCAATTTTATTAATAATTTTACTTTATA 11652
QY 816 TGTTTTTTTCACATTTTGTAAACAAATCATATCATATGATTTGAAAGAGAGGAAATTTGA 875
DB 11651 TGTTTTTTTCACATTTTGTAAACAAATCATATCATATGATTTGAAAGAGAGGAAATTTGA 11592
```



Qy 876 CAGTGAGTAAATGATGAGAAAAAATGTTATTTCTTAAAAAACCCTAAACAAAC 935  
Db 11591 CAGTGAGTAAATGATGAGAAAAAATGTTATTTCTTAAAAAACCCTAAACAAAC 11532

Qy 936 ATGTATCTACTCTCTATTTCAATCTATCTCAATTTCAATTTCTCTTTATCTCTCTTT 995  
Db 11531 ATGTATCTACTCTCTATTTCAATCTATCTCTCAATTTCAATTTCTCTTTATCTCTCTTT 11472

Qy 996 ATTTTATATCATATTTTCAATTAATTTTATTTTACTCTCTTTATTTTCTCTCTAT 1055  
Db 11471 ATTTTATATCATATTTTCAATTAATTTTATTTTACTCTCTTTATTTTCTCTCTAT 11412

Qy 1056 CCCTCTCTATTTCCACTCATATATACACTCCAAAATGGGGCATGCTTTATCACTACT 1115  
Db 11411 CCCTCTCTATTTCCACTCATATATACACTCCAAAATGGGGCATGCTTTATCACTACT 11352

Qy 1116 CTATCTCTCTCACTAAATCAATTAATGAACCTGAAAGCAATGGGCAAGTCTCTCCCT 1175  
Db 11351 CTATCTCTCTCACTAAATCAATTTAAATGAAACTGAAAGCAATGGGCAAGTCTCTCCCT 11292

Qy 1176 CCTCAAGTGATTTCCAACTCAGCAATGCGCATCTGATTTGATTTGATTTGATTTGATTT 1235  
Db 11291 CCTCAAGTGATTTCCAACTCAGCAATGCGCATCTAATTTGATTTGATTTGATTTGATTT 11232

Qy 1236 GTAAAAGTCTTTCCCAATACATACTAATTAATTAATCTTAAATATAAATAAGGATAAAAT 1295  
Db 11231 GTAAAAGTCTTTCCCAATACATACTAATTAATTAATCTTAAATATAAATAAGGATAAAAT 11172

Qy 1296 ATTTTATTTTCTTCATAAAAATTAATAATGATTTATTTTGTGTAGTGATATTTGGAATA 1355  
Db 11171 ATTTTATTTTCTTCATAAAAATTAATAATGATTTATTTTGTGTAGTGATATTTGGAATA 11112

Qy 1356 AATCTAAATATGATATGATTTTATTTTATTTGATTTAAACATATAATCAATATTAATAT 1415  
Db 11111 AATCTAAATATGATATGATTTTATTTTATTTGATTTAAACATATAATCAATATTAATAT 11052

Qy 1416 GATATTTTATATAGTTGTGACACATAATTTTATTAAGGATAAAAAATATGATAAAAT 1475  
Db 11051 GATATTTTATATAGTTGTGACACATAATTTTATTAAGGATAAAAAATATGATAAAAT 10992

Qy 1476 AAATTTTAAATATTTTATTTTATTTACGAGAAAAAATAATTTAGCCATAAATAATGAC 1535  
Db 10991 AAATTTTAAATATTTTATTTTATTTACGAGAAAAAATAATTTAGCCATAAATAATGAC 10932

Qy 1536 CAGCATATTTTACAACTTAGTAATTCATAATTCCTATATGATATTTGCAATTTAAAAA 1595  
Db 10931 CAGCATATTTTACAACTTAGTAATTCATAATTCCTATATGATATTTGCAATTTAAAAA 10872

Qy 1596 CAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTGCCATTTGTTTTCATGCAAA 1655  
Db 10871 CAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTGCCATTTGTTTTCATGCAAA 10812

Qy 1656 CAGAAAGGACGAAAAACCACTCAGATGATCCTCTTCAACCAATTTTACTAGCAA 1715  
Db 10811 CAGAAAGGACGAAAAACCACTCAGATGATCCTCTTCAACCAATTTTACTAGCAA 10752

Qy 1716 ACAAGTCTCAACACTCAAGCCAGCTCTCTCCGTTTCTTTTACACACTTTCTTTGA 1775  
Db 10751 ACAAGTCTCAACACTCAAGCCAGCTCTCTCTCCGTTTCTTTTACACACTTTCTTTGA 10692

Qy 1776 AATAGTAGTATTTTATTTTTCACATGATTTTATTAACGTGCCCCAAGATGCTTTTGAATAG 1835  
Db 10691 AATAGTAGTATTTTATTTTTCACATGATTTTATTAACGTGCCCCAAGATGCTTTTGAATAG 10632

Qy 1836 AGTGCAATTTGTAATGTAATCTAATTAAGAAACATGAAAAAGCATTTGTTCTAACACGATA 1895  
Db 10631 AGTGCAATTTGTAATGTAATCTAATTAAGAAACATGAAAAAGCATTTGTTCTAACACGATA 10572

Qy 1896 ATCTGTGAAGGCGTTAACTCCAAAGATCCAAATTTCACTATATAAATTTGTGACGAAGCA 1955  
Db 10571 ATCTGTGAAGGCGTTAACTCCAAAGATCCAAATTTCACTATATAAATTTGTGACGAAGCA 10512

Qy 1956 AAATGAATTCACATAGCTGAGAGAGAAAGAAAGTTAACTAAGAGCAATACTTTCA 2012  
Db 10511 AAATGAATTCACATAGCTGAGAGAGAAAGAAAGTTAACTAAGAGCAATACTTTCA 10455

RESULT 7  
US-10-776-889-13  
; Sequence 13, Application US/10776889  
; Publication No. US20040158052A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Zhan-Bin  
; APPLICANT: Kinney, Anthony  
; TITLE OF INVENTION: Seed Specific Promoters  
; FILE REFERENCE: BB1531 US NA  
; CURRENT APPLICATION NUMBER: US/10/776,889  
; CURRENT FILING DATE: 2004-02-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 13  
; LENGTH: 1883  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-776-889-13

Query Match 93.6%; Score 1883; DB 19; Length 1883;  
Best Local Similarity 100.0%; Pred. No. 2.6e-226;  
Matches 1883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 GATATTTTTCGCTTCTTCAATATCTGATTTTACCAAAAAAAGCTGTATTTTCTCTAC 189  
Db 1 GATATTTTTCGCTTCTTCAATATCTGATTTTACCAAAAAAAGCTGTATTTTCTCTAC 60

Qy 190 ACTCTCAAGCTTTGTTTTCGCTTCTGACTCTCATGATTTTCTTATATGCAATCACTCT 249  
Db 61 ACTCTCAAGCTTTGTTTTCGCTTCTGACTCTCATGATTTTCTTATATGCAATCACTCT 120

Qy 250 ATTTATAAATGGCATAGGTAGTGAAACAATTTGCAAGCTTGTCAATCAAAAGCTTGCAA 309  
Db 121 ATTTATAAATGGCATAGGTAGTGAAACAATTTGCAAGCTTGTCAATCAAAAGCTTGCAA 180

Qy 310 TGTACAAATTAATGTTTTCATGCCCTTCAAAATATCTGCAACCTAGTCAATTAATCT 369  
Db 181 TGTACAAATTAATGTTTTCATGCCCTTCAAAATATCTGCAACCTAGTCAATTAATCT 240

Qy 370 AACATCTAAGTAGGCTAGTGAATTTTTCGAATAGTCAATGCAAGTCAATTAATCT 429  
Db 241 AACATCTAAGTAGGCTAGTGAATTTTTCGAATAGTCAATGCAAGTCAATTAATCT 300

Qy 430 TGAATTTTGGCTTTGACTCCAAACACTGGCCCGTACATCCGTCCTCAATACATGAAA 489  
Db 301 TGAATTTTGGCTTTGACTCCAAACACTGGCCCGTACATCCGTCCTCAATACATGAAA 360

Qy 490 AGAAATTTGTTTATATTTCTTAATTAATAATATGTCCTTCTAAATTTTCAATATAGTTA 549  
Db 361 AGAAATTTGTTTATATTTCTTAATTAATAATATGTCCTTCTAAATTTTCAATATAGTTA 420

Qy 550 ATTTATATATTTCTTTTCTATTTCTATTTAGTCTATTTTCAAAATATTTATTTATGCA 609  
Db 421 ATTTATATATTTCTTTTCTATTTCTATTTAGTCTATTTTCAAAATATTTATTTATGCA 480

Qy 610 TATGTAAGTACATATATATTTTGTCTATATACATAAATATTTCTAAATATTTAAAAAAG 669  
Db 481 TATGTAAGTACATATATATTTTGTCTATATACATAAATATTTCTAAATATTTAAAAAAG 540

Qy 670 ACTGATATGAAATTTATTTTAAAGCTATATCATTTTATATATATTTTCTTTTCTTTT 729  
Db 541 ACTGATATGAAATTTATTTTAAAGCTATATCATTTTATATATATTTTCTTTTCTTTT 600

Qy 730 CTTTCTTTTCAATTTCTATTTCAATTTAAAGAAATTTTGTAAATTTTATTTATTTATC 789  
Db 601 CTTTCTTTTCAATTTCTATTTCAATTTAAAGAAATTTTGTAAATTTTATTTATTTATC 660

Qy 790 AATTTATAAAAAATTTTACTTTTATATGTTTTTTTCAATTTTTTGTAAACAAATCATATC 849



```
Db 661 AATTTTATAAAATATTTTACATTTATATGTTTTTACACATTTTGTGTTAAACAAATCATATC 720
Qy 850 ATTATGATTTGAAAGAGAGAAATTTGACAGTGAGTAATAAGTGATGAGAAAAAATGTGTT 909
Db 721 ATTATGATTTGAAAGAGAGAAATTTGACAGTGAGTAATAAGTGATGAGAAAAAATGTGTT 780
Qy 910 ATTTTCTTAAAAAACCTTAAACAAACATGTATCTACTCTCTATTTTCATCTATCTCTCATTT 969
Db 781 ATTTTCTTAAAAAACCTTAAACAAACATGTATCTACTCTCTATTTTCATCTATCTCTCATTT 840
Qy 970 TCATTTTCTCTTATCTCTTCTTTTATTTTATCATATCATATCATTTTCCACATTAATTTT 1029
Db 841 TCATTTTCTCTTATCTCTTCTTTTATTTTATCAATCATTTCCATTAATTTT 900
Qy 1030 TTACTCTCTTTTATTTTCTCTATCCCTCTCTTATTTTCCACTCATATATACACTCCAA 1089
Db 901 TTACTCTCTTTATTTTCTCTATCCCTCTCTTATTTTCCACTCATATATACACTCCAA 960
Qy 1090 AATTGGGGCATGCCCTTTATCACTACTCTATCTCTCTCGACTAAATCATTTAAATGAAACTG 1149
Db 961 AATTGGGGCATGCCCTTTATCACTACTCTATCTCTCTCCACTAAATCATTTAAATGAAACTG 1020
Qy 1150 AAAAGCATTTGGCAAGTCTCTCCCTCCCTCAAGTATTTCCAACTCAGCATTTGGCATCTG 1209
Db 1021 AAAAGCATTTGGCAAGTCTCTCCCTCCCTCAAGTATTTCCAACTCAGCATTTGGCATCTG 1080
Qy 1210 ATTGATTCAGTATATCTATTTGCATGTGTAAAGTCTTTTCCAAATACATACTATTAATT 1269
Db 1081 ATTGATTCAGTATATCTATTTGCATGTGTAAAGTCTTTTCCAAATACATACTATTAATT 1140
Qy 1270 AATCTTAAATAAATAGGATAAAATATTTTTTTTTTCTTCAATAAATAAATATGTTAT 1329
Db 1141 AATCTTAAATAAATAGGATAAAATATTTTTTTTTTCTTCAATAAATAAATATGTTAT 1200
Qy 1330 TTTTGTGTTAGATGATATTCGAATAATCTAAATATATGATATGATATGATTTTATATGTA 1389
Db 1201 TTTTGTGTTAGATGATATTCGAATAATCTAAATATATGATATGATATGATTTTATATGTA 1260
Qy 1390 TTAAACATATAATCAATATTAATATGATATTTTTTTTATATAGGTGTGACACATAATTTT 1449
Db 1261 TTAAACATATAATCAATATTAATATGATATTTTTTTTATATAGGTGTGACACATAATTTT 1320
Qy 1450 ATAAGGATAAAAAATATGATAAAATAAATTTTAAATATTTTTTATATTTTACGAGAAAAA 1509
Db 1321 ATAAGGATAAAAAATATGATAAAATAAATTTTAAATATTTTTTATATTTTACGAGAAAAA 1380
Qy 1510 AAATATTTTAGCCATAAATAATGACCAGCATATTTTACACCTTAGTATTCATTAATTT 1569
Db 1381 AAATATTTTAGCCATAAATAATGACCAGCATATTTTACACCTTAGTATTCATTAATTT 1440
Qy 1570 CCTATATGATATTTGAAATTTAAAAACAGATAATCGTTTAAAGGAAAGGAATCCTACGTAT 1629
Db 1441 CCTATATGATATTTGAAATTTAAAAACAGATAATCGTTTAAAGGAAAGGAATCCTACGTAT 1500
Qy 1630 CTCTTGCCATTTGTTTTTATGCAAAACAGAAAGGAGGAGAAAAACCACTCACCATGAATC 1689
Db 1501 CTCTTGCCATTTGTTTTTATGCAAAACAGAAAGGAGGAGAAAAACCACTCACCATGAATC 1560
Qy 1690 ACTCTTCACACCATTTTATGCAAAACAGAAAGGAGGAGAAAAACCACTCACCATGAATC 1749
Db 1561 ACTCTTCACACCATTTTATGCAAAACAGAAAGGAGGAGAAAAACCACTCACCATGAATC 1620
Qy 1750 GTTTCCTTTTACAACACTTTTCTTTGAAATAGTAGTATTTTTTTTTTTCACATGATTTATTA 1809
Db 1621 GTTTCCTTTTACAACACTTTTCTTTGAAATAGTAGTATTTTTTTTTTTCACATGATTTATTA 1680
Qy 1810 CGTCCCAAAAGATGCTTTATGAAATAGAGTGACATTTTGTAAATGACTAAATTTAGAAC 1869
Db 1681 CGTCCCAAAAGATGCTTTATGAAATAGAGTGACATTTTGTAAATGACTAAATTTAGAAC 1740
Qy 1870 TGAANAAGCATTTGTTTAAACAGTAAATCTCTGTGAAGCGTTTAACTCCAAAGATCCAAAT 1929
```

```
Db 1741 TGAANAAGCATTTGTTTAAACAGTAAATCTCTGTGAAGCGTTTAACTCCAAAGATCCAAAT 1800
Qy 1930 TCATATATATAAATTTGTGACGAAAGCAAAATGAATTCATAGCTGAGAGAGAAAGGAAAG 1989
Db 1801 TCATATATATAAATTTGTGACGAAAGCAAAATGAATTCATAGCTGAGAGAGAAAGGAAAG 1860
Qy 1990 GTTAACTTAAGAGCAATACTTCA 2012
Db 1861 GTTAACTTAAGAGCAATACTTCA 1883

RESULT 8
US-10-776-889-14
; Sequence 14, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Zhan-Bin
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Seed Specific Promoters
; FILE REFERENCE: BBI531 US NA
; CURRENT APPLICATION NUMBER: US/10/776,889
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Glycine max
US-10-776-889-14

Query Match 85.4%; Score 1719; DB 19; Length 1719;
Best Local Similarity 100.0%; Pred. No. 8.1e-206;
Matches 1719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 294 CATCAAAAGCTGCAATGTACAAATTAATGTTTTCATGCTTTCAAAATTTATCTGCACC 353
Db 1 CATCAAAAGCTGCAATGTACAAATTAATGTTTTCATGCTTTCAAAATTTATCTGCACC 60
Qy 354 CCCTAGCTATTAATCTAAACATCTAAGTAAGGCTAGTGAATTTTTCGAATAGTCAATGAG 413
Db 61 CCCTAGCTATTAATCTAAACATCTAAGTAAGGCTAGTGAATTTTTCGAATAGTCAATGAG 120
Qy 414 TGCAATTAATTTCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCGGTACATCCGT 473
Db 121 TGCAATTAATTTCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCGGTACATCCGT 180
Qy 474 CCCTCATTTACATGAAAGAAATATTTGTTATATTTCTTAATTTAAATAATTTGTCCTTTCTA 533
Db 181 CCCTCATTTACATGAAAGAAATATTTGTTATATTTCTTAATTTAAATAATTTGTCCTTTCTA 240
Qy 534 AATTTTCATATAGTTAATTAATATATTTTCTTCTTATTTCTTATTTAGTTCTATTTTCA 593
Db 241 AATTTTCATATAGTTAATTAATATATTTTCTTCTTATTTCTTATTTAGTTCTATTTTCA 300
Qy 594 AATTAATTTATGTCATATGTAAGTACATATATTTTTCCTATATATTTTCAATATTTCT 653
Db 301 AATTAATTTATGTCATATGTAAGTACATATATTTTTCCTATATATTTTCAATATTTCT 360
Qy 654 AAATTTATTAATAAAGACTGATATGAAATAATTTTCTTTTAAAGCTATATCATCTTTAT 713
Db 361 AAATTTATTAATAAAGACTGATATGAAATAATTTTCTTTTAAAGCTATATCATCTTTAT 420
Qy 714 ATATACCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 773
Db 421 ATATACCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 480
Qy 774 TAAATTTTATTTATCAATTTTATTAATAAATATTTTCTTTTATATGTTTTTTCACATTTT 833
Db 481 TAAATTTTATTTATCAATTTTATTAATAAATATTTTCTTTTATATGTTTTTTCACATTTT 540
Qy 834 TTAACAACAAATCATATCATATGATGTAAGAGAGGAAATTTGACAGTGAGTAATAAGTGAT 893
Db 541 TTAACAACAAATCATATCATATGATGTAAGAGAGGAAATTTGACAGTGAGTAATAAGTGAT 600
```

QY 894 GAGAAAAAATGCTTATTTCTTAAACCAACATGATGATCTCTCTATT 953  
DB |||||||  
QY 954 TCATCTATCTCATTTTCATTTTCTCTTATCTCTTCTCTTCTTTTATCATATCAT 1013  
DB |||||||  
QY 661 TCATCTATCTCTCATTTTCATTTTCTCTTATCTCTTCTCTTCTTTTATCATATCAT 720  
DB |||||||  
QY 1014 TTCACATTAATATTTTATCTCTCTTATTTTCTCTCTATCCCTCTCTTATTCACCT 1073  
DB |||||||  
QY 721 TTCACATTAATATTTTATCTCTTATTTTCTCTCTATCCCTCTCTTATTCACCT 780  
DB |||||||  
QY 1074 CATATATACACTCCAAATTTGGGCGATGCGCTTTATCACTACTCTATCTCTCCCACTAAAT 1133  
DB |||||||  
QY 781 CATATATACACTCCAAATTTGGGCGATGCGCTTTATCACTACTCTATCTCTCCCACTAAAT 840  
DB |||||||  
QY 1134 CATTAAATGAAACTGAAAGCATTTGGCAAGTCTCTCTCCCTCTCTCAAGTGATTTCCAAC 1193  
DB |||||||  
QY 841 CATTAAATGAAACTGAAAGCATTTGGCAAGTCTCTCTCCCTCTCTCAAGTGATTTCCAAC 900  
DB |||||||  
QY 1194 TCAGCATTTGGCATCTGATTCAGTATCATTTGCGATGCTGTAAGGCTTTTCCACAA 1253  
DB |||||||  
QY 901 TCAGCATTTGGCATCTGATTCAGTATCATTTGCGATGCTGTAAGGCTTTTCCACAA 960  
DB |||||||  
QY 1254 TACATAACTATTAATTAATCTTAAATAAATAAGGATAAAATATTTTTTTTCTTCATAA 1313  
DB |||||||  
QY 961 TACATAACTATTAATTAATCTTAAATAAATAAGGATAAAATATTTTTTTTCTTCATAA 1020  
DB |||||||  
QY 1314 AATTAAATATGTTATTTTGTGTTAGATGTAATTCGAATAAATCTAAATATATGATAA 1373  
DB |||||||  
QY 1021 AATTAAATATGTTATTTTGTGTTAGATGTAATTCGAATAAATCTAAATATATGATAA 1080  
DB |||||||  
QY 1374 TGATTTTATATGATTAACATATATCAATTAATTAATATGATATTTTTTTTATATAGG 1433  
DB |||||||  
QY 1081 TGATTTTATATGATTAACATATATCAATTAATTAATATGATATTTTTTTTATATAGG 1140  
DB |||||||  
QY 1434 TTGTACACATAAATTTTAAAGGATAAATAATGATAAATAAATTTTAAATATTTTAA 1493  
DB |||||||  
QY 1141 TTGTACACATAAATTTTAAAGGATAAATAATGATAAATAAATTTTAAATATTTTAA 1200  
DB |||||||  
QY 1494 TATTTAGGAGAAAAAATATTTTAGCCATAAATAAATGACAGCATATTTTACACCT 1553  
DB |||||||  
QY 1201 TATTTAGGAGAAAAAATATTTTAGCCATAAATAAATGACAGCATATTTTACACCT 1260  
DB |||||||  
QY 1554 TAGTAATTCATAATCCCTATATGATATTTGAAATTTAAACACAGATAATCGTTAAGGGA 1613  
DB |||||||  
QY 1261 TAGTAATTCATAATCCCTATATGATATTTGAAATTTAAACACAGATAATCGTTAAGGGA 1320  
DB |||||||  
QY 1614 AGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAAACAGAAAGGACGAAAAAC 1673  
DB |||||||  
QY 1321 AGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAAACAGAAAGGACGAAAAAC 1380  
DB |||||||  
QY 1674 CACCTCAACATGAATCACTCTTCACACCATTTTACAGCAAAACAGTCTCAACACTGA 1733  
DB |||||||  
QY 1381 CACCTCAACATGAATCACTCTTCACACCATTTTACTAGCAAAACAGTCTCAACAACTGA 1440  
DB |||||||  
QY 1734 AGCAGCTCTCTTCCGTTTCTTTTACACACCTTTCTTTTGAATAGTAGTATTTTTTTT 1793  
DB |||||||  
QY 1441 AGCAGCTCTCTTCCGTTTCTTTTACACACCTTTCTTTTGAATAGTAGTATTTTTTTT 1500  
DB |||||||  
QY 1794 TCACATGATTTATTAACGTCGCAAAAGATGCTTATTTGAATAGAGTGACATTTTGAATGT 1853  
DB |||||||  
QY 1501 TCACATGATTTATTAACGTCGCAAAAGATGCTTATTTGAATAGAGTGACATTTTGAATGT 1560  
DB |||||||  
QY 1854 ACTACTAATTTAGAACATGAAAAAGCATTTGTTCTAAACAGATAATTCCTGTGAAGCGGTTAA 1913  
DB |||||||  
QY 1561 ACTACTAATTTAGAACATGAAAAAGCATTTGTTCTAAACAGATAATTCCTGTGAAGCGGTTAA 1620  
DB |||||||  
QY 1914 CTCACAGATGCCAATTTCACTATATAAATTTGTCGAAAGCAAAATGAATTTACATAGCT 1973  
DB |||||||  
QY 1621 CTCACAGATGCCAATTTCACTATATAAATTTGTCGAAAGCAAAATGAATTTACATAGCT 1680  
DB |||||||

QY 1974 GAGAGAGAAAGGAAGGTTAACTTAAGAACCAATACTTCA 2012  
DB |||||||  
QY 1681 GAGAGAGAAAGGAAGGTTAACTTAAGAACCAATACTTCA 1719  
DB |||||||  
RESULT 9  
US-10-776-889-15  
; Sequence 15, Application US/10776889  
; Publication No. US20040158052A1  
; GENERAL INFORMATION:  
; APPLICANT: Lib, Zhan-Bin  
; APPLICANT: Kinney, Anthony  
; TITLE OF INVENTION: Seed Specific Promoters  
; FILE REFERENCE: BB1531 US NA  
; CURRENT APPLICATION NUMBER: US/10/776,889  
; CURRENT FILING DATE: 2004-02-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 1553  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-776-889-15  
Query Match 77.2%; Score 1553; DB 19; Length 1553;  
Best Local Similarity 100.0%; Pred. No. 4.4e-185;  
Matches 1553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 460 CCCCGTACATCCGTCCTCATACATGAAAAGAAATATTGTTTATATTCTTAATTAATAA 519  
DB |||||||  
QY 520 TATTGTCCTTCTAAATTTTCATATAGTTAAATTTATTTATATTACTTTTTCTCTATTCTAT 579  
DB |||||||  
QY 61 TATTGTCCTTCTAAATTTTCATATAGTTAAATTTATTTATATTACTTTTTCTCTATTCTAT 120  
DB |||||||  
QY 580 TAGTCTATTTTCAAAATTTATTTATGCATATGTAAGTACATTAATTTTTCGTATAT 639  
DB |||||||  
QY 121 TAGTCTATTTTCAAAATTTATTTATGCATATGTAAGTACATTAATTTTTCGTATAT 180  
DB |||||||  
QY 640 ACTTAATATTTTCTAAATTTTAAAAAAGACGTGATATGAAAATTTATCTTTTAAAG 699  
DB |||||||  
QY 181 ACTTAATATTTTCTAAATTTTAAAAAAGACGTGATATGAAAATTTATCTTTTAAAG 240  
DB |||||||  
QY 700 CTATATCATTTTATATATATCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 759  
DB |||||||  
QY 241 CTATATCATTTTATATATATCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 300  
DB |||||||  
QY 760 AGAAATAAATTTTGTAAATTTTATTTATCAATTTTATAAAATATTTTACTTTTATGTT 819  
DB |||||||  
QY 301 AGAAATAAATTTTGTAAATTTTATTTATCAATTTTATAAAATATTTTACTTTTATGTT 360  
DB |||||||  
QY 820 TTTTTCATATTTTGTAAACCAATCATATCATTTGAAAGAGAGAGGAAATTTGACAGT 879  
DB |||||||  
QY 361 TTTTTCATATTTTGTAAACCAATCATATCATTTGAAAGAGAGAGGAAATTTGACAGT 420  
DB |||||||  
QY 880 GAGTAATAAGTATGAGAAAAAAATGTTATTTTCTTAAAAAACCTTAAACCAACATGT 939  
DB |||||||  
QY 421 GAGTAATAAGTATGAGAAAAAAATGTTATTTTCTTAAAAAACCTTAAACCAACATGT 480  
DB |||||||  
QY 940 ATCTACTCTATTTTCAATCTATCTCTCATTTTCAATTTTCTTTTATCTCTTTCTTTATTT 999  
DB |||||||  
QY 481 ATCTACTCTATTTTCAATCTATCTCTCATTTTCAATTTTCTTTTATCTCTTTCTTTATTT 540  
DB |||||||  
QY 1000 TTTTATCATATCATTTTCAATTAATTTTACTCTCTTTATTTTCTCTCTATCCCT 1059  
DB |||||||  
QY 541 TTTTATCATATCATTTTCAATTAATTTTACTCTCTTTATTTTCTCTCTATCCCT 600  
DB |||||||  
QY 1060 CTCCTATTTTCCACTCATATATACACTCCAAAATTTGGGCGATGCTTTTATCAGTACTCTAT 1119  
DB |||||||  
QY 601 CTCCTATTTTCCACTCATATATACACTCCAAAATTTGGGCGATGCTTTTATCAGTACTCTAT 660  
DB |||||||  
QY 1120 CTCCTCCACTAAATCATTTTAAATGAAACCTGAAAAGCATTTGGCAAGTCTCTCTCCCTCCTC 1179  
DB |||||||

Db 661 CTTCCCTCCACTAAATCATTTTAAATGAAACGAAAGGCAATTTGGCAAGTCTCTCCCTCCCTC 720  
Qy 1180 AAGTGATTTTCCAACTCAGCAATTTGGCATCTGATTTGATTCAGTATATCTATTGTCATGTGTA 1239  
Db 721 AAGTGATTTTCCAACTCAGCAATTTGGCATCTGATTTGATTCAGTATATCTATTGTCATGTGTA 780  
Qy 1240 AAGTCTTTCCCAACAATACATCACTATTAAATCTTAAATATAAATAAAGGATAAAATATTT 1299  
Db 781 AAGTCTTTCCCAACAATACATCACTATTAAATCTTAAATATAAATAAAGGATAAAATATTT 840  
Qy 1300 TTTTCTTTCTCATAAATTAATAATATGTTATTTTGTGTTAGATGTATATTCGAATAATC 1359  
Db 841 TTTTCTTTCTCATAAATTAATAATATGTTATTTTGTGTTAGATGTATATTCGAATAATC 900  
Qy 1360 TAAATATATGATAATGATTTTATATGATTAATAAATCAATATCAATATTAATATGATA 1419  
Db 901 TAAATATATGATAATGATTTTATATGATTAATAAATCAATATCAATATTAATATGATA 960  
Qy 1420 TTTTCTTTATATAGTTGTACACATATTTTATAAGGATAAATAATATGATAAATAAT 1479  
Db 961 TTTTCTTTATATAGTTGTACACATATTTTATAAGGATAAATAATATGATAAATAAT 1020  
Qy 1480 TTTTAAATATTTTATATTTACGAGAAATAAATAATTTTAGCCATAAATATGACAGC 1539  
Db 1021 TTTTAAATATTTTATATTTACGAGAAATAAATAATTTTAGCCATAAATATGACAGC 1080  
Qy 1540 ATATTTTACAACTTAGTAATTTTATATATTTTCCATATGATATTTGAAATTAATAAACA 1599  
Db 1081 ATATTTTACAACTTAGTAATTTTATATATTTTCCATATGATATTTGAAATTAATAAACA 1140  
Qy 1600 TAACTGTTAAGGAAAGGAATCTAGCTCATCTCTTGCCATTTGTTTTCATGCAAAACA 1659  
Db 1141 TAACTGTTAAGGAAAGGAATCTAGCTCATCTCTTGCCATTTGTTTTCATGCAAAACA 1200  
Qy 1660 AAGGACGAAATAACCACTCACCATGATCTCTTCCACCAATTTTACTAGCAAAACA 1719  
Db 1201 AAGGACGAAATAACCACTCACCATGATCTCTTCCGTTCTTTTACACACATTTCTTTGA 1260  
Qy 1720 GTCTCAACCACTGAAGCAGCTCTCTTCCGTTCTTTTACACACATTTCTTTGA 1779  
Db 1261 GTCTCAACCACTGAAGCAGCTCTCTTCCGTTCTTTTACACACATTTCTTTGA 1320  
Qy 1780 GTAGTATTTTTCATCATGATTTTAAACGTGCAAAAGATGCTTATGAAATAGAGTG 1839  
Db 1321 GTAGTATTTTTCATCATGATTTTAAACGTGCAAAAGATGCTTATGAAATAGAGTG 1380  
Qy 1840 CACATTTGTAATGTACTACTAATTAAGACATGAAAGCATTTGTTCTAACACGATAATCC 1899  
Db 1381 CACATTTGTAATGTACTACTAATTAAGACATGAAAGCATTTGTTCTAACACGATAATCC 1440  
Qy 1900 TGTGAAGCGTTAACTCCAAAGATCCAATTTTCACTATATAAATTTGTGACGAAAGCAAAAT 1959  
Db 1441 TGTGAAGCGTTAACTCCAAAGATCCAATTTTCACTATATAAATTTGTGACGAAAGCAAAAT 1500  
Qy 1960 GAAATTCATAGCTGAGAGAGAAAGGATTAACGTAAGAGCAATATCTTCA 2012  
Db 1501 GAAATTCATAGCTGAGAGAGAAAGGATTAACGTAAGAGCAATATCTTCA 1553

RESULT 10

US-10-776-889-16  
; Sequence 16, Application US/10776889  
; Publication No. US20040158052A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Zhan-Bin  
; APPLICANT: Kinney, Anthony  
; TITLE OF INVENTION: Seed Specific Promoters  
; FILE REFERENCE: BBI531 US NA  
; CURRENT APPLICATION NUMBER: US/10/776, 889  
; CURRENT FILING DATE: 2004-02-11  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97

; SEQ ID NO 16  
; LENGTH: 1367  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-776-889-16

Query Match 67.9%; Score 1367; DB 19; Length 1367;  
Best Local Similarity 100.0%; Pred. No. 7.5e-162;  
Matches 1367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 646 ATATTTCTTAAATTTTAAAGAAAGACTGATATGAAATTTTATCTCTTTTAAAGCTATAT 705  
Db 1 ATATTTCTTAAATTTTAAAGAAAGACTGATATGAAATTTTATCTCTTTTAAAGCTATAT 60  
Qy 706 CATTTTATATATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 765  
Db 61 CATTTTATATATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 120  
Qy 766 AAATTTTGTAAATTTTATTTTATCAATTTTATAAATAATTTTATCTTTTATATGTTTTTCA 825  
Db 121 AAATTTTGTAAATTTTATTTTATCAATTTTATAAATAATTTTATCTTTTATATGTTTTTCA 180  
Qy 826 CATTTTGTAAACAAATCATATCATTTATGATTTGAAAGAGAGAAATTTGACAGTGAGTAA 885  
Db 181 CATTTTGTAAACAAATCATATCATTTATGATTTGAAAGAGAGAAATTTGACAGTGAGTAA 240  
Qy 886 TAAGTGATGAGAAAAAATGTGTTATTTCTTAAAAAAAACCTTAAACAAATGATGATCTAC 945  
Db 241 TAAGTGATGAGAAAAAATGTGTTATTTCTTAAAAAAAACCTTAAACAAATGATGATCTAC 300  
Qy 946 TCTCTATTTTCACTCTCTCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1005  
Db 301 TCTCTATTTTCACTCTCTCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 360  
Qy 1006 CATATCATTTTCAATTAATTTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1065  
Db 361 CATATCATTTTCAATTAATTTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 420  
Qy 1066 TTTCCACTCATATATACACTCCAAATTTGGGGCATGCTTTTATCACTACTCTATCTCTCTC 1125  
Db 421 TTTCCACTCATATATACACTCCAAATTTGGGGCATGCTTTTATCACTACTCTATCTCTCTC 480  
Qy 1126 CACTAAATCATTTAAATGAAACTGAAAGCAATTTGGCAAGTCTCTCTCTCTCTCTCAAGTGA 1185  
Db 481 CACTAAATCATTTAAATGAAACTGAAAGCAATTTGGCAAGTCTCTCTCTCTCTCTCAAGTGA 540  
Qy 1186 TTTCCAACTCAGCAATTTGGCATCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1245  
Db 541 TTTCCAACTCAGCAATTTGGCATCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 600  
Qy 1246 TTCCCAATACATATACTATTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTT 1305  
Db 601 TTCCCAATACATATACTATTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTT 660  
Qy 1306 CTTCAATAAATTTAAATATGTTATTTTGTGTTAGATGTATATTTTCGAATAAATCTAAATA 1365  
Db 661 CTTCAATAAATTTAAATATGTTATTTTGTGTTAGATGTATATTTTCGAATAAATCTAAATA 720  
Qy 1366 TATGATATGATTTTATTTTATTTGATTTAAACATATAATCAATTTAAATTTAAATTTAAATTTT 1425  
Db 721 TATGATATGATTTTATTTTATTTGATTTAAACATATAATCAATTTAAATTTAAATTTAAATTTT 780  
Qy 1426 TATATAGTTGTACACATAAATTTTATAAGGATAAATAATGATAAATAAATTTAAATTTTAA 1485  
Db 781 TATATAGTTGTACACATAAATTTTATAAGGATAAATAATGATAAATAAATTTAAATTTTAA 840  
Qy 1486 TATTTTATATTTACGAGAAAAAATAATTTTATAGCCATAAATAAATGACAGCATATTT 1545  
Db 841 TATTTTATATTTACGAGAAAAAATAATTTTATAGCCATAAATAAATGACAGCATATTT 900  
Qy 1546 TACAACCTTAGTAATTCATTAATTTCTTATATGTTATTTTGAATTTTAAATTTAAATTTTAAATTCG 1605  
Db 901 TACAACCTTAGTAATTCATTAATTTCTTATATGTTATTTTGAATTTTAAATTTAAATTTTAAATTCG 960

Qy	1606	TTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTTGTGTTTTCATGCAAAACAGAAAGGGA	1665
Db	961	TTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTTGTGTTTTCATGCAAAACAGAAAGGGA	1020
Qy	1666	CGAAAAACCACTCACCATGAATCAGTTCACACCACTTTTACTAGCAAAACAGTCTCA	1725
Db	1021	CGAAAAACCACTCACCATGAATCAGTTCACACCACTTTTACTAGCAAAACAGTCTCA	1080
Qy	1726	ACAACTGAAGCCAGCTCTCTTCCGTTCTTTTACAACTTTCTTTGAAATAGTAGTA	1785
Db	1081	ACAACTGAAGCCAGCTCTCTTCCGTTCTTTTACAACTTTCTTTGAAATAGTAGTA	1140
Qy	1786	TTTTTTTTTTCACATGATTTATTAAAGTGCAGAAAGATGCTTTATTGAATAGAGTGCACATT	1845
Db	1141	TTTTTTTTTTCACATGATTTATTAAAGTGCAGAAAGATGCTTTATTGAATAGAGTGCACATT	1200
Qy	1846	TGTAATGTACTACTAATTTAGAACATGAATAAGCAATGTTCTTAACGATAGTGCACATT	1905
Db	1201	TGTAATGTACTACTAATTTAGAACATGAATAAGCAATGTTCTTAACGATAGTGCACATT	1260
Qy	1906	GGCGTTAACTCCAAAGATCCAAATTTCACTATATAAATTTGTGACGAAAGCAAAATGAATTC	1965
Db	1261	GGCGTTAACTCCAAAGATCCAAATTTCACTATATAAATTTGTGACGAAAGCAAAATGAATTC	1320
Qy	1966	ACATAGCTGAGAGAGAAAGGAAAGGTTAACTTAAGAAAGCAATACTTCA	2012
Db	1321	ACATAGCTGAGAGAGAAAGGAAAGGTTAACTTAAGAAAGCAATACTTCA	1367
RESULT 11			
US-10-776-889-17			
; Sequence 17, Application US/10776889			
; Publication No. US20040158052A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Zhan-Bin			
; TITLE OF INVENTION: Seed Specific Promoters			
; FILE REFERENCE: BB1531 US NA			
; CURRENT APPLICATION NUMBER: US/10/776,889			
; CURRENT FILING DATE: 2004-02-11			
; NUMBER OF SEQ ID NOS: 22			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 17			
; LENGTH: 1160			
; TYPE: DNA			
; ORGANISM: Glycine max			
US-10-776-889-17			
Query Match			
Best Local Similarity 100.0%; Score 1160; DB 19; Length 1160;			
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	853	ATGATTGAAAGAGAGGAAATTTGACAGTGAAGTAATAGTGAATAGAGAGAAATGTTGTTATT	912
Db	1	ATGATTGAAAGAGAGGAAATTTGACAGTGAAGTAATAGTGAATAGAGAGAAATGTTGTTATT	60
Qy	913	TCCTAAAAAACCTAAACAAACATGTACTCTCTATTTTCATCTATCTCTCATTTTCA	972
Db	61	TCCTAAAAAACCTAAACAAACATGTACTCTCTATTTTCATCTATCTCTCATTTTCA	120
Qy	973	TTTTTCTCTTTATCTCTTTTCTTTTATTTTATCATATCATTTTACATTAATTTATTTTA	1032
Db	121	TTTTTCTCTTTATCTCTTTTCTTTTATTTTATCATATCATTTTACATTAATTTATTTTA	180
Qy	1033	CTCTCTTTATTTTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAT	1092
Db	181	CTCTCTTTATTTTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAT	240
Qy	1093	TGGGGCATGCTTTATCAGTACTCTATCTCTCCCACTAAATCATTTAAATGAAACTGAAA	1152
Db	241	TGGGGCATGCTTTATCAGTACTCTATCTCTCCCACTAAATCATTTAAATGAAACTGAAA	300

Qy	1153	AGCATTTGGAAGTCTCTCCCTCCTCAAGTGAATTTTCCAACTCAGCATTTGCGCATCTGATT	1212
Db	301	AGCATTTGGAAGTCTCTCCCTCCTCAAGTGAATTTTCCAACTCAGCATTTGCGCATCTGATT	360
Qy	1213	GATTTCAGTATATCTATTGATGATGTAAGTCTTTCCCAATACATACTATTAAATTAAT	1272
Db	361	GATTTCAGTATATCTATTGATGATGTAAGTCTTTCCCAATACATACTATTAAATTAAT	420
Qy	1273	CTTTAAATAAATAAGGATAAAATATTTTCTTTTCTTATATAAATTAATATGCTTATTTT	1332
Db	421	CTTTAAATAAATAAGGATAAAATATTTTCTTTTCTTATATAAATTAATATGCTTATTTT	480
Qy	1333	TTGTTTAGATGTATTTTCGAATAAACTTAATAATATGATAATGATTTTATATGATTA	1392
Db	481	TTGTTTAGATGTATTTTCGAATAAACTTAATAATATGATAATGATTTTATATGATTA	540
Qy	1393	ACATATATATCAATATTAATATGATATTTTATATATAGGTTGTACACATAATTTTATA	1452
Db	541	ACATATATATCAATATTAATATGATATTTTATATATAGGTTGTACACATAATTTTATA	600
Qy	1453	AGGATAAAAAATATGATAAAAAATAAATTTTAAATATTTTATATTTACGAGAAAAA	1512
Db	601	AGGATAAAAAATATGATAAAAAATAAATTTTAAATATTTTATATTTACGAGAAAAA	660
Qy	1513	TATTTTAGCCATAAAATAATGACAGCATATTTTACAACCTTAGTAATTTCAATAATTCCT	1572
Db	661	TATTTTAGCCATAAAATAATGACAGCATATTTTACAACCTTAGTAATTTCAATAATTCCT	720
Qy	1573	ATATGTATATTTGAAATTAATAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTC	1632
Db	721	ATATGTATATTTGAAATTAATAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTC	780
Qy	1633	TTGCCATTTGTTTTCATGCAACAGAAAGGAGGAGAAACCAACCTCACCATGAATCACT	1692
Db	781	TTGCCATTTGTTTTCATGCAACAGAAAGGAGGAGAAACCAACCTCACCATGAATCACT	840
Qy	1693	CTTCACACCATTTTACTAGCAAAACAGTCTCAACAACTGAAGCCAGCTCTCTTCCGTT	1752
Db	841	CTTCACACCATTTTACTAGCAAAACAGTCTCAACAACTGAAGCCAGCTCTCTTCCGTT	900
Qy	1753	TCCTTTTACACACATTTCTTGAATAGTAGTATTTTTCACATGATTTTATTAACGT	1812
Db	901	TCCTTTTACACACATTTCTTGAATAGTAGTATTTTTCACATGATTTTATTAACGT	960
Qy	1813	GCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGA	1872
Db	961	GCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGA	1020
Qy	1873	AAAGCATTTGTTCTAAACAGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAAATTTCA	1932
Db	1021	AAAGCATTTGTTCTAAACAGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAAATTTCA	1080
Qy	1933	CTATATAAATTTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGATTT	1992
Db	1081	CTATATAAATTTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGATTT	1140
Qy	1993	AACTAAGAACCAATACTTCA	2012
Db	1141	AACTAAGAACCAATACTTCA	1160
RESULT 12			
US-10-776-889-18			
; Sequence 18, Application US/10776889			
; Publication No. US20040158052A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Zhan-Bin			
; TITLE OF INVENTION: Seed Specific Promoters			
; FILE REFERENCE: BB1531 US NA			
; CURRENT APPLICATION NUMBER: US/10/776,889			
; CURRENT FILING DATE: 2004-02-11			
; NUMBER OF SEQ ID NOS: 22			

; SOFTWARE: Microsoft Office 97									
; SEQ ID NO 18									
; LENGTH: 967									
; TYPE: DNA									
; ORGANISM: Glycine max									
US-10-776-889-18									
Query Match 48.1%; Score 967; DB 19; Length 967;									
Best Local Similarity 100.0%; Pred. No. 6.8e-112;									
Matches 967; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1046	TTCTCTCTATCCCTCTCTATTTTCACATCATATATACACTCCAAATTTGGGCGATGCCTT	1105						
Db	1	TTCTCTCTATCCCTCTCTATTTCCACATCATATATACACTCCAAATTTGGGCGATGCCTT	60						
Qy	1106	TATCACTACTCTATCTCTCCACTAAATCATTTAAATGAAACTGAAAAAGCATTTGGCAAGT	1165						
Db	61	TATCACTACTCTATCTCTCCACTAAATCATTTAAATGAAACTGAAAAAGCATTTGGCAAGT	120						
Qy	1166	CTCTCCCTCTCTCAAGTGATTTTCCAACTCAGCATTTGGCATCTGATTCAGTATATC	1225						
Db	121	CTCTCCCTCTCTCAAGTGATTTTCCAACTCAGCATTTGGCATCTGATTCAGTATATC	180						
Qy	1226	TATTGCATGTGTAAAGTCTTTCCACAATACATACTAATTAATCTTAAATAAATAA	1285						
Db	181	TATTGCATGTGTAAAGTCTTTCCACAATACATACTAATTAATCTTAAATAAATAA	240						
Qy	1286	AGGATAAAATATTTTTTTTCTTCATAAAATTTAAATATGTATTTTTTTTGTAGATGTA	1345						
Db	241	AGGATAAAATATTTTTTTTCTTCATAAAATTTAAATATGTATTTTTTTTGTAGATGTA	300						
Qy	1346	TATTCGAATAAATCTAAATATATGATATGATTTTTTTTATATGATTAACATATAATCAA	1405						
Db	301	TATTCGAATAAATCTAAATATATGATATGATTTTTTTTATATGATTAACATATAATCAA	360						
Qy	1406	TATTAAATATCATATTTTTTTTATATAGTCTGACACATAATTTTATAGGATAAATAA	1465						
Db	361	TATTAAATATCATATTTTTTTTATATAGTCTGACACATAATTTTATAGGATAAATAA	420						
Qy	1466	TGATAAAATATAAATTTTAAATATTTTTTATATTTACGAGAAAAAATAATTTTAGCCATA	1525						
Db	421	TGATAAAATATAAATTTTAAATATTTTTTATATTTACGAGAAAAAATAATTTTAGCCATA	480						
Qy	1526	AATAAATGACGAGCATATTTTCAACCTTAGTAATTTCAATAATTCCTATATGTATTTG	1585						
Db	481	AATAAATGACGAGCATATTTTCAACCTTAGTAATTTCAATAATTCCTATATGTATTTG	540						
Qy	1586	AAATTTAAAAACAGATAATCGTTTACGAGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTT	1645						
Db	541	AAATTTAAAAACAGATAATCGTTTACGAGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTT	600						
Qy	1646	TTTCATGCAAAACAGAAAGGAGGAAAAACCACTCACCATGAATCACTCTTCACACCATTTT	1705						
Db	601	TTTCATGCAAAACAGAAAGGAGGAAAAACCACTCACCATGAATCACTCTTCACACCATTTT	660						
Qy	1706	TTATAGCAAAACAAGTCTCAACAAGTGAAGCGAGCTCTCTTTCCGTTTCTTTTACACA	1765						
Db	661	TTATAGCAAAACAAGTCTCAACAAGTGAAGCGAGCTCTCTTTCCGTTTCTTTTACACA	720						
Qy	1766	CTTTCTTTGAATAGTAGTATTTTTTTTTCACATGATTTTATTAACGTCGCAAAAGATGCT	1825						
Db	721	CTTTCTTTGAATAGTAGTATTTTTTTTTCACATGATTTTATTAACGTCGCAAAAGATGCT	780						
Qy	1826	TATTGAATAGAGTGACATTTTGTATGTACTACTAAATTAGAACATGAAAAAGCATTTGTC	1885						
Db	781	TATTGAATAGAGTGACATTTTGTATGTACTACTAAATTAGAACATGAAAAAGCATTTGTC	840						
Qy	1886	TAAACAGATAATCTCTGTAAGCGGTTAACTCCAAAGATCCAAATTTCACTATATAAATGT	1945						
Db	841	TAAACAGATAATCTCTGTAAGCGGTTAACTCCAAAGATCCAAATTTCACTATATAAATGT	900						
Qy	1946	GACGAAAGCAAAATGAATTCATAGCTGAGAGAGAAAGGAAAGTTTAACCTAAGAGCAA	2005						
; Query Match 38.3%; Score 770; DB 19; Length 770;									
Best Local Similarity 100.0%; Pred. No. 2.7e-87;									
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1243	TCITTCCACAATACATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1302						
Db	1	TCITTCCACAATACATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	60						
Qy	1303	TTTCTTCATAAAATTAATAATATGTTATTTTTTTTGTAGATGATATTCGAATAAATCTAA	1362						
Db	61	TTTCTTCATAAAATTAATAATATGTTATTTTTTTTGTAGATGATATTCGAATAAATCTAA	120						
Qy	1363	ATATATGATATGATTTTTTTTATATTTGATTAACAATATAANTCAATTAATTAATTAAT	1422						
Db	121	ATATATGATATGATTTTTTTTATATTTGATTAACAATATAANTCAATTAATTAAT	180						
Qy	1423	TTTTATATAGTTTGTACACATAAATTTTATAGGATAAATAATATGATAAATAAATTTT	1482						
Db	181	TTTTATATAGTTTGTACACATAAATTTTATAGGATAAATAATATGATAAATAAATTTT	240						
Qy	1483	AAATATTTTTTATATTTACGAGAAAAAATAATTTTACGACATAAATAATGACCAAGATA	1542						
Db	241	AAATATTTTTTATATTTACGAGAAAAAATAATTTTACGACATAAATAATGACCAAGATA	300						
Qy	1543	TTTTACAACTTAGTAAATTCATAAATTCCTATATGTATATTTTGAATTTAAAAACAGATA	1602						
Db	301	TTTTACAACTTAGTAAATTCATAAATTCCTATATGTATATTTGAATTTAAAAACAGATA	360						
Qy	1603	TCGTTAAGGAAGGAATTCCTACGTCATCTCTGTCATCTTTGTTTTCATGCAACAGAAAG	1662						
Db	361	TCGTTAAGGAAGGAATTCCTACGTCATCTCTGTCATCTTTGTTTTCATGCAACAGAAAG	420						
Qy	1663	GGAGAAAAACCACTCACCATGAATCAGTCCTTTCACACCATTTTTTCTAGCAACAAGTC	1722						
Db	421	GGAGAAAAACCACTCACCATGAATCAGTCCTTTCACACCATTTTTTCTAGCAACAAGTC	480						
Qy	1723	TCACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTTACAACACTTTCTTTGAAATAGTA	1782						
Db	481	TCACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTTACAACACTTTCTTTGAAATAGTA	540						
Qy	1783	GTATTTTTTTTTCACATGATTTTATTAACGTCGCAAAAGATGCTTATTGAATAGAGTGCAC	1842						
Db	541	GTATTTTTTTTTCACATGATTTTATTAACGTCGCAAAAGATGCTTATTGAATAGAGTGCAC	600						
Qy	1843	ATTGTGTAATCTACTTAATTAAGACATGAAAGGCAATTTGTTCTTAACACCATATCTCTGT	1902						
Db	601	ATTGTGTAATCTACTTAATTAAGACATGAAAGGCAATTTGTTCTTAACACCATATCTCTGT	660						

```
QY 1903 GAAGCGTTAACTCAAAGATCCAAATTTCACTATATAAATTTGTGACGAAAGCAAAATGAA 1962
Db 661 GAAGCGTTAACTCAAAGATCCAAATTTCACTATATAAATTTGTGACGAAAGCAAAATGAA 720
QY 1963 TTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTTAAGAACCAATACTTCA 2012
Db 721 TTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTTAAGAACCAATACTTCA 770

RESULT 14
US-10-776-889-20
; Sequence 20, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Zhan-Bin
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Seed Specific Promoters
; FILE REFERENCE: BB1531 US NA
; CURRENT APPLICATION NUMBER: US/10/776,889
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Glycine max
US-10-776-889-20

Query Match 29.0%; Score 584; DB 19; Length 584;
Best Local Similarity 100.0%; Pred. No. 4.5e-64;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 ATAGGTTGTACACATAATTTTATAAGGATAAAATATGATAAAATATGATAAAATATTAATAT 1488
Db 1 ATAGGTTGTACACATAATTTTATAAGGATAAAATATGATAAAATATGATAAAATATTAATAT 60
QY 1489 TTTTATATTTACGAGAAAAAATAATTTTAGCCATAAATAAATGACGACATATTTTAC 1548
Db 61 TTTTATATTTACGAGAAAAAATAATTTTAGCCATAAATAAATGACGACATATTTTAC 120
QY 1549 AACCTTAGTAATTCATAAATTCCTATATGATATATTTGAAATTTAAATAACAGATAATCGTTA 1608
Db 121 AACCTTAGTAATTCATAAATTCCTATATGATATATTTGAAATTTAAATAACAGATAATCGTTA 180
QY 1609 AGGAAGGAATCCTACGTCATCTCTGCCATTTGTTTTTCATGCAACACAGAAAGGACGA 1668
Db 181 AGGAAGGAATCCTACGTCATCTCTGCCATTTGTTTTTCATGCAACACAGAAAGGACGA 240
QY 1669 AAAACCCACCTCACCATGAATCACTCTTTCACACCAATTTTACTAGCAAAACAAGTCTCAACA 1728
Db 241 AAAACCCACCTCACCATGAATCACTCTTTCACACCAATTTTACTAGCAAAACAAGTCTCAACA 300
QY 1729 ACTGAAGCCAGCTCTCTTTCGGTTTCTTTTAAACAATTTCTTTGAAATAGTAGTATTT 1788
Db 301 ACTGAAGCCAGCTCTCTTTCGGTTTCTTTTAAACAATTTCTTTGAAATAGTAGTATTT 360
QY 1789 TTTTTCACATGATTTTAACTGCGCCAAAGATGCTTATTTGAATAGAGTGCAATTTGT 1848
Db 361 TTTTTCACATGATTTTAACTGCGCCAAAGATGCTTATTTGAATAGAGTGCAATTTGT 420
QY 1849 AATGTACTACTAATTAGAACATGAAAAAGCATTTGTTCTTAACACGATAATCCTGTGAAGGC 1908
Db 421 AATGTACTACTAATTAGAACATGAAAAAGCATTTGTTCTTAACACGATAATCCTGTGAAGGC 480
QY 1909 GTTAACTCCAAAGATCCAAATTTCACTATATAAATTTGTGACGAAAGCAAAATGAATTCACA 1968
Db 481 GTTAACTCCAAAGATCCAAATTTCACTATATAAATTTGTGACGAAAGCAAAATGAATTCACA 540
QY 1969 TAGCTGAGAGAGAAAGGAAAGGTTAACTTAAGAACCAATACTTCA 2012
Db 541 TAGCTGAGAGAGAAAGGAAAGGTTAACTTAAGAACCAATACTTCA 584
```

```
RESULT 15
US-10-776-889-21
; Sequence 21, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Zhan-Bin
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Seed Specific Promoters
; FILE REFERENCE: BB1531 US NA
; CURRENT APPLICATION NUMBER: US/10/776,889
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Glycine max
US-10-776-889-21

Query Match 21.1%; Score 425; DB 19; Length 425;
Best Local Similarity 100.0%; Pred. No. 3.1e-44;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1588 ATTAAAAACAGATAATCGTTAAGGGAAAGAAATCCTACGTCATCTCTTGCCATTTGTTTT 1647
Db 1 ATTAAAAACAGATAATCGTTAAGGGAAAGAAATCCTACGTCATCTCTTGCCATTTGTTTT 60
QY 1648 CATGCAAAACAGAAAGGACGAAAAACCCACTACCATGAATCACTCTTTCACACCATTTTT 1707
Db 61 CATGCAAAACAGAAAGGACGAAAAACCCACTCACCATGAATCACTCTTTCACACCATTTTT 120
QY 1708 ACTAGCAAAACAAGTCTCAACAACCTGAAGCCAGCTCTCTTTCGGTTTCTTTTACAAACACT 1767
Db 121 ACTAGCAAAACAAGTCTCAACAACCTGAAGCCAGCTCTCTTTCGGTTTCTTTTACAAACACT 180
QY 1768 TTCTTTGAAATAGTAGTATTTTTCACATGATTTTAAACGTGCCAAAAGATGCTTA 1827
Db 181 TTCTTTGAAATAGTAGTATTTTTCACATGATTTTAAACGTGCCAAAAGATGCTTA 240
QY 1828 TTGAATAGAGTGACACATTTTGTAATGTACTACTAATTTAGAACATGAAAAAGCATTTGTTCTA 1887
Db 241 TTGAATAGAGTGACACATTTTGTAATGTACTACTAATTTAGAACATGAAAAAGCATTTGTTCTA 300
QY 1888 ACACGATAATCCTGTGAAGCGTTAACTCCAAAGATCCAAATTTCACTATATAAATTTGTGA 1947
Db 301 ACACGATAATCCTGTGAAGCGTTAACTCCAAAGATCCAAATTTCACTATATAAATTTGTGA 360
QY 1948 CGAAAGCAAAATGAATTCATAGCTGAGAGAGAAAGGAAAGGTTAACTTAAGAACCAATA 2007
Db 361 CGAAAGCAAAATGAATTCATAGCTGAGAGAGAAAGGAAAGGTTAACTTAAGAACCAATA 420
QY 2008 CTTCA 2012
Db 421 CTTCA 425
```

Search completed: September 2, 2005, 16:33:38  
Job time : 989.534 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 22:19:32 ; Search time 6107.49 Seconds  
(without alignments)  
15962.684 Million cell updates/sec

Title: US-10-776-889-1

Perfect score: 2012

Sequence: 1 atctaggcccttgattata.....aactaagaagcaataacttca 2012

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_ars.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	170.2	8.5	111861	9	AC069435 Homo sapi
c 2	170	8.4	349980	6	AX344555 Sequence
c 3	168.2	8.4	90550	9	AL592166 Human DNA
c 4	162	8.1	179668	2	CR626936 Danio rer
c 5	161.2	8.0	14867	3	AE001398 Plasmidii
c 6	160.4	8.0	104992	2	AC005504 Plasmidii
c 7	160.4	8.0	169546	2	AC004157 Plasmidii
c 8	160.4	8.0	202169	2	AC149549 Gorilla g
c 9	160.4	8.0	250421	3	AE014849 Plasmidii
c 10	159.8	7.9	180096	2	CR626872 Danio rer
c 11	159.6	7.9	178427	2	CR392006 Danio rer
c 12	158.6	7.9	120029	2	HSJ282H10
c 13	158.4	7.9	143331	9	AC091214 Homo sapi
c 14	158.4	7.9	226108	2	CR759967 Danio rer
c 15	158	7.9	129240	9	AC084128 Homo sapi
c 16	157.8	7.8	170627	2	AC125567 Rattus no
c 17	157.8	7.8	170627	2	AC125567 Rattus no
c 18	157.6	7.8	181792	9	AC098822 Homo sapi
c 19	156.8	7.8	125836	2	AC137629 Homo sapi
c 20	156.8	7.8	79018	9	AL627107 Human DNA

20	156.8	7.8	135121	9	AC069525	Homo sapi
21	156	7.8	15548	6	AX347057	Sequence
c 22	155.2	7.7	148937	2	CR361568	Danio rer
c 23	155.2	7.7	348174	3	CR382399	Plasmidii
c 24	154.6	7.7	172777	2	CR847809	Danio rer
c 25	154.6	7.7	176704	2	CR628364	Danio rer
c 26	154.4	7.7	170143	2	AC048384	Homo sapi
c 27	154.2	7.7	171537	2	CR762390	Danio rer
c 28	154.2	7.7	174198	2	CR762438	Danio rer
c 29	154	7.7	169800	2	CR847821	Danio rer
c 30	154	7.7	254050	3	PFA929358	Plasmidii
c 31	154	7.7	313050	3	PFA929352	Plasmidii
c 32	153.6	7.6	186165	9	AC112187	Homo sapi
c 33	153.4	7.6	159618	9	AC096750	Homo sapi
c 34	153.2	7.6	157467	2	CR792423	Danio rer
c 35	153	7.6	149563	2	CR376750	Danio rer
c 36	152.8	7.6	154044	2	CR786570	Danio rer
c 37	152.8	7.6	214709	2	CR749183	Danio rer
c 38	152.8	7.6	228431	2	CR788237	Danio rer
c 39	152.6	7.6	210218	2	CR555304	Danio rer
c 40	152.6	7.6	210871	2	CR759970	Danio rer
c 41	152.4	7.6	148075	9	AC110014	Homo sapi
c 42	152.4	7.6	155106	9	AC104069	Homo sapi
c 43	151.8	7.5	251237	2	CR812792	Danio rer
c 44	151.6	7.5	178785	9	AC068139	Homo sapi
c 45	151.6	7.5	227073	2	CR846086	Danio rer

## ALIGNMENTS

RESULT 1  
AC069435/c  
LOCUS  
DEFINITION Homo sapiens 3 BAC RP11-384L1 (Roswell park Cancer Institute Human BAC Library) complete sequence.  
ACCESSION AC069435  
VERSION AC069435.19  
KEYWORDS GI:20340410  
SOURCE HTG.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 111861)  
Muzny,D.M., Adams,C., Ali-Osman,B., Ali-Osman,F.R., Allen,C., Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbacia,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulesged,H., Lozada,R.J., Lu,X., Lucier,A., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N.,

AC069435 111861 bp DNA linear PRI 30-AUG-2002  
Homo sapiens 3 BAC RP11-384L1 (Roswell park Cancer Institute Human BAC Library) complete sequence.

AC069435 GI:20340410

HTG.

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 111861)

Muzny,D.M., Adams,C., Ali-Osman,B., Ali-Osman,F.R., Allen,C.,

Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,

Barbacia,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,

Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,

Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,

Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,

Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,

Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,

Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,

Loulesged,H., Lozada,R.J., Lu,X., Lucier,A., Luna,R.,

Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,

Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,

Mel,G., Metzker,M., Miner,Z., Mitchell,T., Mohabbat,K.,

Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,

Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,

Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N.,





```
/rpt_family="GA-rich"
20595..20717
/rpt_family="FLAM_C"
20721..20763
/rpt_family=" (TAAA)n"
21442..21725
/rpt_family="MLT1A1"
complement(21762..22107)
/rpt_family="MLT11"
22360..22395
/rpt_family="AT-rich"
22559..22817
/rpt_family="AluSx"
complement(22927..23015)

Query Match      8.5%; Score 170.2; DB 9; Length 111861;
Best Local Similarity 48.6%; Pred. No. 5,1e-11;
Matches 551; Conservative 0; Mismatches 578; Indels 4; Gaps 3;

Qy 465 TACATCCGCTCAATTACATGAAAGAAATATCTTTATATCTTAAATTAATAATATG 524
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
44345 TATATTTATATATTTATATATACACACCACTTTGTTTATATATAATATATATAAAT 44286

Qy 525 TCCCTCTCAAAATTTTCATATAGTTAAATTTATATATCTTTTCTCTATTTAGTT 584
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
44285 ATATATATTTTATATATATATATAATATATATATAATAATATATATATTTTATATATTT 44226

Qy 585 CTATTTTCAAAATTTATTTATGCAATGTAAAGTACATTTATTTTTCGTATATACTTA 644
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
44225 ATATATAATATTTATATATATATTTTATATATTTTATATATAATATATATATATAGTA 44166

Qy 645 AATATTTCTAAATTTAAAAAGACGATATGAAAAATTTATTTCTTTTAAAGCTATA 704
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
44165 AATATATTTTATATATTTATATATTTTATATTTTATATATATATATATATATATATTT 44106

Qy 705 TCAITTTATATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 764
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
44105 TTTATATATTTATATATATATTTTATATATATTTTATATATATTTTATATATTTTATATTTA 44046

Qy 765 TAAATTTTGTAAATTTTATTTATCAATTTTATAAAATTTTATCTTTTATATGTTTTTTC 824
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
44045 TATATATTTTATATATTTTATATATATTTTATATATATTTTATATATATATATATTTAT 43986

Qy 825 ACATTTTGTAAACAAATCATATCATTTATGATGAAAGAGAGGAAATGACAGTGAGTA 884
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43985 ATATTATATTTATATATATTTATATATTTTATATATATTTATATATATATATATATTTTA 43926

Qy 885 ATAAGTGATGAGAAAAATGTTTATTTCTTAAAAAAACCTTAAACAAACATGTATCTA 944
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43925 TATATATTTATATATATTTTATATATATTTTATATATATTTATATATATTTATATATATA 43866

Qy 945 CTCCTATTTTCATCTATCTCTCATTTCTATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1004
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43865 TTATATATTTTATATATATTTATATATATTTTATATATATTTATATATATTTATATATA 43806

Qy 1005 TCATATCATTTTCACATTAATTTTATCTC-TCTTTATTTTCTCTCTATCCCTCTCT 1063
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43805 TATTTATATATTTATATATATTTATATATATTTTATATATATTTTATATATATATATATA 43746

Qy 1064 TATTTCCACTCATATATACATCCCAAAATTTGGGGCATGCGCTTTATCATCTACTCTATCTCC 1123
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43745 CATTTTATATATTTTATATATATTTTATATATATTTTATATATATTTTATATATATATAT 43686

Qy 1124 TCCACTCAATCAATTTTAAATGAAACGTGAAAGCATTTGGCAAGTCTCTCCCTCTCAAGT 1183
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43685 ATTTATATATATAGATATATATATATATATTTTATATATATTTATATATATTTATATATAT 43626

Qy 1184 GATTTCCAACTCAGCATTTGGCATCTGATTTGATTCAGTATATCTATGCTATGTAAGTA 1243
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43625 ATTTTATATTTATATATATATATATTTTATATATATTTTATATATATTTTATATATATAT 43566

Qy 1244 CTTTCCCAATACATAACTTAAATCTTAAATTAATAAAGATATAAATATTTTTTTT 1303
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 43565 ATTATATATATTT-ATATATATTTTATATATATTTTATATATATTTTATATATTTTAAATATATTTAT 43507
Qy 1304 TTCTTCATAAAATTAATAATATGTTTATTTTGTGTTTAGATGTATATTCGAATAAAATCTAAA 1363
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43506 ATATTATATATATATATATATTTTATATATATTTTATATATATTTATATATATTTATATATTTA 43447

Qy 1364 TATATGATAATGATTTTATATATTTGATTAACATATATAATCAATATTAATTAATATGATATTTT 1423
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43446 TATATATTTATATATATATATATATTTTATATATTTTATATATTTTATATATATTTATATATAT 43387

Qy 1424 TTTATATAGGTTGTAC--ACATAATTTTATAAGGATAAAAAATATGATAAAAAATAAAATTT 1481
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43386 AATATATAATATATATTTTATATATATTTTATATATATTTTATATATATTTTATATATATTTATAT 43327

Qy 1482 TAAATATTTTATATATTTTACGAGAAAAAAAATATTTTAGCCATATAAATAATGACCAGCAT 1541
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43326 TATATATATTTTATATATATTTTATATATTTTATATATATTTTATATATATTTTATATATATA 43267

Qy 1542 ATTTTACACCTTAGTAAATTCATAAAATTCCTATATGATATATTTTGAAATTAATAA 1594
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43266 TTTTATATGTATAAAATTTATATATATTTTATATATAAAAAATATATAAATTTATATA 43214

RESULT 2
AX344555
LOCUS AX344555 349980 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 6 from Patent WO0200932.
ACCESSION AX344555
VERSION AX344555.1 GI:18492441
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of known genetic parameters within the mhc
JOURNAL Patent: WO 0200932-A 6 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
source 1..349980
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)-Original length of seq 1: 3,673,778 <223>-splitted as follows.-seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001 649.980-seq 03 600.001 949.980-seq 04 900.001 1.249.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001 1.849.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001 2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001 3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001 3.649.980-seq 13 3.600.001 3.673.778 <223>-Original length of seq 2: 3,673,778 <223>-splitted as follows.-seq 14 0.000.001 TO 0.349.980-seq 15 0.300.001 649.980-seq 16 600.001 949.980-seq 17 900.001 1.249.980-seq 18 1.200.001 1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001 2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001 2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001 3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001 3.673.778"
```

## ORIGIN

```
Query Match      8.4%; Score 170; DB 6; Length 349980;
Best Local Similarity 48.8%; Pred. No. 3,6e-11;
Matches 543; Conservative 0; Mismatches 565; Indels 4; Gaps 3;

Qy 490 AGAAATATGTTTATATCTTAAATAAAATATGTCCTCTCTAAATTTTCATATAGTTA 549
Db 213576 ATATATTTTATATATATTTTATATATTTTATATATATTTTATATATATTTTATATATTTT 213635

Qy 550 ATTATATTATATACTTTTTCTCTATTTCTATTTAGTTCTTCAAAATTTATTTATGCA 609
Db 213636 ATATATATTTTATATATTTTATATATATTTTATATATATTTTATATATTTTATATATA 213695
```





	Matches	498;	Conservative	0;	Mismatches	605;	Indels	3;	Gaps	37
Qy	498	TGTTATATCTTAAATTAAAAAATATTGTCCCTCTCAAATTTTCATATAGTAATTAATTAT	557							
Dd	73226	TA	73167							
Qy	558	ATTACITTTTTCTCTACTTCTATTAGTTCCTATTTTCAAAATTTATTTATGCAATATGAAGA	617							
Dd	73166	TA	73107							
Qy	618	GTACATTTATTTTTTGCTATATATCTTAAATATTTCTAAAATTTATTA AAAAAGACGTGATAT	677							
Dd	73106	TATATATATATATAT - AT	73048							
Qy	678	GA AAAAATTTTATCTTTTTTAAAGCTATATCATTTTTTATATATATCTTTTTCTTTTTCTTTCT	737							
Dd	73047	AT	72988							
Qy	738	TCATTTTCTATTCAAATTTAAATAGAAATAAAT - TTTGTAAAATTTTATTTATCACAATTTAT	796							
Dd	72987	AT	72928							
Qy	797	AAAAATATTTTACTTTTATATATGTTTTTTCATTTTTTGGTTTAAACCAATCATATCATATATGA	856							
Dd	72927	ATCAT	72868							
Qy	857	TTGAAAGAGAGGAAATTTGCAGTGCAGTAAAGTGATGAGAAAAAAATCGTTATTTTCTCT	916							
Dd	72867	AT	72808							
Qy	917	AAAAAAACCTAAACAAACATGATCTACTCTCTATTTCCACTCATATATACACTCCA AAATTTGGG	1096							
Dd	72807	TA	72748							
Qy	977	TCCTCTTATCTCTCTTTTATTTTTTTTATCATATCATTTTCACATTAATTTATTTTACTCT	1036							
Dd	72747	TA	72688							
Qy	1037	CTTTATTTTTTCTCTCTATCCCCTCTCTTATTTCCACTCATATATACACTCCA AAATTTGGG	1096							
Dd	72687	TGTATATGTATATAAT	72628							
Qy	1097	GCATGCCCTTTATCACACTCTATCTCCTCCACTPAAATCATTTTAAATGAAACTGAA AAGCA	1156							
Dd	72627	ATATATATACNN	72568							
Qy	1157	TTGGCAAGCTCTCCCTCCCTCCTCAAGTGATTTCCAACTCAGCAATGGCACTGTATGTATT	1216							
Dd	72567	NN	72508							
Qy	1217	CAGTATATCTATTGCGATGCTGTA AAAAGCTTTCCACAAATACATAACTTATTAATTAATCTTA	1276							
Dd	72507	ATATATATATATAAATTTA	72448							
Qy	1277	AATAAAATAAAGGATAAAAATTTTTTTTTTCTTCATATA - AAAATAAAAATATGTTATTTTTTG	1335							
Dd	72447	TA	72388							
Qy	1336	TTTAGATGTATATTCGAATAAAATCTAAAATATATATATATATATATATATATATATATATAT	1395							
Dd	72387	TA	72328							
Qy	1396	ATATAATCATATATAAATATGATATTTTTTTTATATAGGTTGTACACATAATTTTATAGG	1455							
Dd	72327	TATATATATATATTTTA	72268							
Qy	1456	ATAAAAAATATGATAAAAAATAAATTTTTTAAATATTTTTTATATTTTTTACGAGAAAAAATAAT	1515							
Dd	72267	TA	72208							
Qy	1516	TTTAGCCATAAATAAATACCAGCAGCATATTTTACAACTTTAGTAATTCATAAAATCCTATA	1575							
Dd	72207	TA	72148							

Qy	1576	TGATATATTTCGAATTAACAGATA	1601
Db	72147	TATATATATATATATATATATATA	72122
RESULT 5			
LOCUS	AE001398	14867 bp DNA linear	INV 04-OCT-2002
DEFINITION	Plasmodium falciparum 3D7 chromosome 2 section 35 of 73 of the complete sequence.		
ACCESSION	AE001398 AE001362		
VERSION	AE001398.1	GI:3845197	
KEYWORDS			
SOURCE	Plasmodium falciparum 3D7		
ORGANISM	Plasmodium falciparum 3D7		
REFERENCE	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
AUTHORS	Gardner,M.J., Tetzelin,H., Carucci,D.J., Cummings,L.M., Aravind,L., Koonin,E.V., Shalloom,S., Mason,T., Yu,K., Fujii,C., Pederson,J., Shen,K., Jing,J., Aston,C., Lai,Z., Schwartz,D.C., Percec,M., Salzberg,S., Zhou,Y., Sutton,G.G., Clayton,R., White,O., Smith,H.O., Fraser,C.M., Adams,M.D., Venter,J.C. and Hoffman,S.L.		
TITLE	Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum		
JOURNAL	Science 282 (5391), 1126-1132 (1998)		
MEDLINE	99021743		
PUBMED	9804551		
REFERENCE	2 (bases 1 to 14867)		
AUTHORS	Gardner M.J., Hall,N., Fung,E., White,O., Bettiman,M., Hyman,R.W., Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.B., Craig,A., Kytes,S., Chan,M.-S., Nene,V., Shalloom,S.J., Suh,B., Peterson,J., Angiuoli,S., Percea,M., Allen,J., Selengut,J., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.		
TITLE	Genome sequence of the human malaria parasite Plasmodium falciparum		
JOURNAL	Nature 419, 498-511 (2002)		
REFERENCE	3 (bases 1 to 14867)		
AUTHORS	Gardner M.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-Oct-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
FEATURES	Location/Qualifiers		
source	1..14867 /organism="Plasmodium falciparum 3D7" /mol_type="genomic DNA" /isolate="3D7" /db_xref="taxon:36329" /chromosome="2" complement (<1..20) /notes="AT rich"		
repeat_region	/rpt_type=tandem complement (65..137) /notes="AT rich"		
repeat_region	/rpt_type=tandem complement (178..211) /rpt_type=tandem /rpt_unit="(ta)n"		
repeat_region	307..327 /rpt_type=tandem complement (328..348) /rpt_type=tandem /rpt_unit="(ta)n"		
repeat_region	378..423 /rpt_type=tandem complement (479..503) /notes="AT rich"		

repeat_region	/rpt_type=tandem complement(513. .539)	repeat_region	/rpt_unit="(ataa)n" 2886. .2924
repeat_region	/rpt_type=tandem /rpt_unit="(a)n" 577. .633	repeat_region	complement(2977. .3167)
repeat_region	/rpt_type=tandem /rpt_unit="(a)n" complement(714. .747)	repeat_region	/rpt_type=tandem /rpt_unit="(taaa)n" 3047. .3132
repeat_region	/rpt_type=tandem /note="AT_rich" 841. .867	repeat_region	/rpt_type=tandem complement(3047. .3132)
repeat_region	/rpt_type=tandem /rpt_unit="(a)n" complement(888. .918)	repeat_region	/rpt_type=tandem /rpt_unit="(ta)n" complement(3258. .3420)
repeat_region	/rpt_type=tandem /note="AT_rich" complement(1006. .1049)	repeat_region	/rpt_type=tandem /note="AT_rich" complement(3450. .3514)
repeat_region	/rpt_type=tandem /note="AT_rich" complement(1075. .1120)	repeat_region	/rpt_type=tandem /rpt_unit="(ta)n" complement(3532. .3559)
repeat_region	/rpt_type=tandem /note="AT_rich" complement(1140. .1283)	repeat_region	/rpt_type=tandem complement(3689. .3720)
repeat_region	/rpt_type=tandem /note="AT_rich" complement(1294. .1346)	repeat_region	/rpt_type=tandem complement(4024. .4114)
repeat_region	/rpt_type=tandem /rpt_unit="(ataa)n" complement(1423. .1531)	repeat_region	/rpt_type=tandem /note="AT_rich" 4107. .4146
gene	/rpt_type=tandem complement(<1570. .>2424)	repeat_region	/rpt_type=tandem complement(4115. .4146)
mRNA	/locus_tag="PFB0490c" complement(<1570. .>2424)	repeat_region	/rpt_type=tandem /rpt_unit="(ta)n" 4227. .4261
CDS	/locus_tag="PFB0490c" complement(1570. .2424)	repeat_region	complement(4348. .4386)
	/product="hypothetical protein" /protein_id="AAC71887.1" /db_xref="GI:3845198"	repeat_region	/rpt_type=tandem /note="AT_rich" complement(4403. .4450)
	/translation="MKENKIMDYLSCLPDDVDREKIKSGKSNLLKSSSTKSKDYKK SSIFSKRDSHKSGSFRGRSGFINRKSQFKPYNNRLINKYNNYKGRNFHGR DNFKRGTSGSFRVFNDRKSGFKKRFISNRNKSYSKSYRNGSNMGRKSNFKAPTR TVWTKRLNYYKTVGAPVKPKNNLISLYRKNRTFALNTKRKSPVGTIKSSVPRKRIK GLKKGSLKSKTRKSTSGSKPKPLNKYFLSKIKIVTSLNKIPSPKLEQKQNVNLPESI NNATTKN"	repeat_region	/rpt_type=tandem complement(5102. .5191)
repeat_region	complement(2150. .2188)	repeat_region	/rpt_type=tandem /note="AT_rich" complement(5246. .5275)
repeat_region	/rpt_type=tandem complement(2446. .2468)	repeat_region	/rpt_type=tandem /note="AT_rich" complement(5437. .5484)
repeat_region	/rpt_type=tandem /rpt_unit="(ta)n" 2504. .2656	repeat_region	/rpt_type=tandem /note="AT_rich" complement(5554. .5577)
repeat_region	/rpt_type=tandem /rpt_unit="(ataa)n" 2684. .2739	repeat_region	/rpt_type=tandem /note="AT_rich" complement(6302. .6350)
repeat_region	/rpt_type=tandem complement(2693. .2736)		
repeat_region	/rpt_type=tandem complement(2697. .2732)		
repeat_region	/rpt_type=tandem /rpt_unit="(ta)n" complement(2757. .2845)		
repeat_region	/rpt_type=tandem /note="AT_rich" complement(2856. .2880)		
repeat_region	/rpt_type=tandem /rpt_unit="(ta)n" 2884. .2953		
repeat_region	/rpt_type=tandem		

Query Match

Best Local Similarity 8.0%; Score 161.2; DB 3; Length 14867;

Mismatches 564; Conservative 0; Mismatches 543; Indels 17; Gaps 6;

Oy

491 GAAATATTGTTATATCTTAATTAATAAAATATTGTCCTTCTAAATTTTCATATAGTTAA 550

|||||

Db

8529 GAAATAATTATTTATATATAATAATAATAATAATAATAATAATAATAATAATAATAATTTTA 8470

|||||

Oy

551 TTATTATATTACTTTTTTCTTCTATTCTATTAGTTCTATTCTTTCAATTTATTATTATGCAT 610

|||||





```
Qy 509 TTAATTAATAAATATTGTCCTCTCTAAATTTTCATATAGTTAATTATATATACTTTTT 568
Db 73671 ATTAATATTAAATATTATGTTATTTTAAATTTGTTA-----ATATTTATATTA 73622

Qy 569 CTCTATTCCTATTAGTCTTATTTTCAAAATTAATTATGTCATATGTAAGTACATTATAT 628
Db 73621 TTTTATTTAAATATATTAAATTTAAATTTAATTATATATATTTAATTATATATAT 73562

Qy 629 TTTTGCTATATACTTAAATATTCTTAAATATTATTAATAAAGACTGATAGAAAATTTAT 688
Db 73561 TTATAATATATTATTAATTTAAATTTGTTATTTTATTTATATATATATATATTTAAT 73502

Qy 689 TCTTTTAAAGCTATATCATTTTATATATACTTTTCTTTCTTTCTTTCTTTCTTTCTAT 748
Db 73501 TATTATTTTATTTATATATATATTTAAATTTTATATATATATATATATATATATAT 73442

Qy 749 TCAATTTAATAAGAAATAAATTTTGTAATTTTATTTATTCATCAATTTATAAAAAATTTTA 808
Db 73441 ATATTTAATTTATTTATATATT-----AATTTAATTTATATTTAATTTAATTTA 73387

Qy 809 CTTTATATGTTTTTTCACATTTTGTGTTAAACAAATCATATCATATGATGAAAGAGAGG 868
Db 73386 TATTTTATTTAATTTATTTTATTTTATTTAATTTAATTTAATTTAATTTAATTTAAT 73327

Qy 869 AAATTCAGCTGAGTAAATAGTGAGAGAAAATGTTATTTCTCTAAAAAAACCTA 928
Db 73326 TAAATTTATATATATTTAATTTAATTTATATATATATATTTAATTTAATTTATATTTA 73267

Qy 929 AACAAACATGATCTCTCTATTTTCATCTCATCTCTCATTTTCTTTCTTTATCTC 988
Db 73266 ATTTATTTATATATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 73207

Qy 989 TTTCTTTATTTTTTATCATATCAFTTCACATTAATTTATTTTACTCTCTTTATTTTTC 1048
Db 73206 TTATAACTATAAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 73147

Qy 1049 TCTCTATCCCTCTTATTTCCACTCATATATACCTCAAAATTTGGGGATGCTTTAT 1108
Db 73146 AATTATATATATTTATGTTTTTAAATTTAATTTAATTTAATTTAATTTAATTTAAT 73087

Qy 1109 CACTACTCTATCTCTCCACTAAATCAFTTTAAATGAAACTGAAAGCAATTTGGCAAGTCTC 1168
Db 73086 TATTATTTAATTAATTTCTTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAT 73027

Qy 1169 CTCCTCCCTCCCAAGTGATTTCCAACTCAGCAATTTGGCAFTCGATTCAGTATATCTAT 1228
Db 73026 TTAATTTATTTAAATATTGTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAT 72967

Qy 1229 TGCATGTGTAAGTCTTTCCACAATACATAAATTTAATTTAATTTAATTTAATTTAAGG 1288
Db 72966 ATATTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 72907

Qy 1289 ATAAAAATTTTTTTTCTTCATAAAAATTTAAAAATATGTTATTTTTT--GTTAGATGTAT 1346
Db 72906 TTATTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 72847

Qy 1347 ATTCGAATAAATCTAAATATATATGATTAATTTTTTTTATATTGATTTAAACATATATCAAT 1406
Db 72846 ATTTAATTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 72787

Qy 1407 ATTTAAATATGATATTTTTTTTATATAGGTTGTACACATAATTTTATAGGATFAAAAAATAT 1466
Db 72786 ATAAATTTATAATTTTTTTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 72727

Qy 1467 GATAAAAATAATTTTAAATATTTTTTATATTTTACGAGAAAAAANAAT---ATTTTAGCCA 1523
Db 72726 TATTATTAATTAATTTATATTTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 72667

Qy 1524 TAAATAAATGACGACATATTTTACAACTTTAGTAATTCATAAAATTCCTATATGATATTT 1583
Db 72666 TATTAAATTTATTTATTTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTA 72607
```

```
Qy 1584 TGAATTAARACAGATAAATCGTTAAGGAGGAATCCTAGTCATCTCTTGCCATTGTT 1643
Db 72606 TATATATTTAATTTTAAATATATATTTTAAATAAATTTTATAATGATTAATTTGATATATTA 72547

Qy 1644 TTTTCATGCAACAGAAA 1661
Db 72546 TATAATTATATATATATA 72529

RESULT 7
AC004157/c
LOCUS
DEFINITION
Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***, 2 unordered pieces.
AC004157
VERSION
HTG; HTGS PHASE1.
KEYWORDS
Plasmodium falciparum (malaria parasite P. falciparum)
SOURCE
Plasmodium falciparum
ORGANISM
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 169546)
Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,
Kurdi,O.B., Conway,A.B. and Davis,R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished
2 (bases 1 to 169546)
Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
Direct Submission
Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
COMMENT
On Aug 12, 2000 this sequence version replaced gi:8810447.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 23466: contig of 23466 bp in length
* 23467 23666: gap of unknown length
* 23667 169546: contig of 145880 bp in length.
FEATURES
Location/Qualifiers
source
1..169546
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"
/clones="PFYAC293"
/clones="3D7"
ORIGIN
Query Match 8.0% Score 160.4; DB 2; Length 169546;
Best Local Similarity 46.6%; Pred. No. 5.5e-10;
Matches 736; Conservative 0; Mismatches 816; Indels 26; Gaps 6;
Qy 95 AGAAAAAAGAAAAAGAGGATCATTTTGAAGGATATTTTTCGCTCTCTATTCAATA 154
Db 85675 ATATTATTAATTAATTTAANAATATAAATTAAGAAGATAATTTTATACATTTTATTATTT 85616

Qy 155 CTGTATTTTACCAAAAAAATGATTTTTCCTACACTCTCAAGCTTGTGTTTTCGGCTTC 214
Db 85615 AAATATATAGTAATAAATAAATTTTATGTTATTTTATTAATAAATATTTTATTATTATT 85556

Qy 215 GACTCTCATGTTCTTCCTCATATGCCAATCACTCTATTAT-----AAATGCCATAAGTA 270
Db 85555 TATTATTTAATAAATAAATTAATTTTATAAATAATATTTTATTTTAAATTAATAATAAAC 85496

Qy 271 GTGTGAACAATTCGAAAGCTTGTTCATCAAAAGCTTGCAATGTACAAAATTAATGTTTTTCA 330
Db 85495 ATATTAATTAATTAATTAATAATATATATATATTTTTTTTAAATAATAATAATAATTTC 85436
```

Qy 331 TGCCTT--TCAAAATATCTGCACCCCTAGCTATTAACTTAACATCTAAGTAAAGCGTAG 388  
Db 85435 CATTTTAAATTAATAAATAAACAATTTATTAACTTAATTAATAATAAATAA 85376  
Qy 389 TGAATTTTTTCGAATAGTCAGAGTCGATTAATTTCCCGTGACTATTTTGGCTTTGAC 448  
Db 85375 ATAAATTTAAATGTGTAAATTAATTAATAATAAACAATTTAATAATAATTAATATATA 85316  
Qy 449 TCCAACACTGGCCCGGTACATCCGCTCATTAACATGAAGAAAGAAATTTGTTTATATTC 508  
Db 85315 ATATTATATATTATCTTAATAATAATAATTTTAAATTAATTTAATAATAATAATAATTT 85256  
Qy 509 TTAATTTAAATAATGTCCCTCTTAATTTTTCATATAGTTAAATTAATTAATTAATCTTTTT 568  
Db 85255 ATTAATTTAAATATGATATTTTATTTAAATGTTTA-----ATAATTAATTA 85206  
Qy 569 CTCTATTTCTATAGTTCTATTTTCCAAATTAATTTATGATGATATGTAAGTACATTAAT 628  
Db 85205 TTTTATTAATATATTAATTAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 85146  
Qy 629 TTTTGCTATATACCTTAATATTTCTAAATTAATTAATAAAGAGACTGATATGAATAATTTAT 688  
Db 85145 TTATAATAATTAATTAATTAATTTATTTATTTATTAATTAATTAATTAATTAATTAAT 85086  
Qy 689 TCTTTTTAAAGCTATATCATTTATATATATCTTTTTCTTTCTTTCTTTCTATTTTCTAT 748  
Db 85085 TATTATTTTATTTATATATATATATATTTAAATTTATATATATATATTTAAATTTATAT 85026  
Qy 749 TCAATTTAAATGAATAATTTTGTAAATTTTATTTATTTATCAATTTATAAATAATTTTA 808  
Db 85025 ATATTTAAATTTATTTATATATTT-----AAATTAATTAATTTATTTAAATTTATA 84971  
Qy 809 CTTTATATGTTTTTTCACATTTTGTAAACAAATCATATCATGATGTAAGAGAGG 868  
Db 84970 TATTTTATTTATTAATTTATTTATTTATTTAAATTTAAATTAATTAATTAATTTAAAT 84911  
Qy 869 AAATGTACAGTGAGTAATAAGTGATGAGAAAAAATGTGTTATTTCTTAAAAAAACCTA 928  
Db 84910 TAAATTAATATATATTTAAATTTAAATTAATATATATTTAAATTTGATATATATTTTA 84851  
Qy 929 AACAAACATGATCTACTCTATTTTCATCTAATCTCATTTTCATTTTCTTTCTTTATCTC 988  
Db 84850 ATTTATTTATATATTTATTTAAATTTATTTATATATTTAAATTTATATATTTTATTTT 84791  
Qy 989 TTTCTTTATTTTATCATATCATTTTCACATTAATTTTCTTTTACTCTCTTTATTTTCTC 1048  
Db 84790 TTATAACTATAAATTAATTAATTTAAATTTAAATTTTATTTTATTTTACTAAATTAAT 84731  
Qy 1049 TCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTTGGGGCATGCCTTAT 1108  
Db 84730 AATTTATATATATTTATTTGTTTAAATTTATTTAAATAATTCATTTTATTTAAATTA 84671  
Qy 1109 CACTACTCTATCTCCTCCACTAAATCATTTAAATGAACATGAAAGCATTTGGCAAGTCTC 1168  
Db 84670 TATTATTTAATAAATTTCTTTAAATTTAATAAATTAATTTGTTTAAATATATATTTATAT 84611  
Qy 1169 CTCCCTCTCAAGTGATTTCCCACTCAGCATTCGCATCTGATGATTTTCAGTATATCTAT 1228  
Db 84610 TTAATTTAATTAATTTGTTTAAATTAATTAATTAATTAATTTATTTATTTAATTTATTT 84551  
Qy 1229 TGCATGTGTAAGTCTTTCCACAATACATAACTATTTAAATTAATTTTAAATAAATAAGG 1288  
Db 84550 ATATTATTTAATTAATTTATATATATTTTATTTATTTTATTTTAAATTTAAATTTAT 84491  
Qy 1289 ATAAATATTTTTTTTCTTCATTAATAATTAATAATGATGTTATTTT--GTTTATGATAT 1346  
Db 84490 TTATTATTTATTTTATTAATTAATTAATTAATTAATTTTATTTTATTTAATTTATTTAA 84431  
Qy 1347 ATTCGAATAATCTAAATATATGATATGATTTTTTATTTGATTTAAACATATAATCAAT 1406  
Db 84430 ATTTAAATTTATTAATTTATATATATTTAAATTAATTTATTAATTTATTTATTTATTT 84371  
Qy 1407 ATTAATATGATATTTTTTTTATATAGTTGTACACATAATTTTATTAAGGAATAAATAAT 1466

Db 84370 ATAAATTAATAATATTTTATTTTATTTTAAATTTTAAATTTGATTAATATTTTCTTATTTAT 84311  
Qy 1467 GATAAAAAATAATTTTAAATATTTTATATTTTACGAGAAAAAATAAT--ATTTTAGCCA 1523  
Db 84310 TATTTATATTAATTAATTCATATTTTATTAATTAATGATATTAATTAATTAATTAATTAATA 84251  
Qy 1524 TAAATAAATGACCCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGATATTT 1583  
Db 84250 TATTTAAATTTATTTATTTTATTTTATTTAATTAATTAATTAATTAATTTATTAATATA 84191  
Qy 1584 TGAATTTAAAAACAGATAATTCGTTAAGGGAAGAAATCCTACGTCATCTCTTGGCAATTTGT 1643  
Db 84190 TATATATTTAAATTTTAAATATATTTTAAATAAATTTTATATATGATTAATTTATATATA 84131  
Qy 1644 TTTTCATGCAACAGAAA 1661  
Db 84130 TATAATTTATATATATATA 84113

RESULT 8  
AC149549/cLOCUS AC149549 202169 bp DNA linear HTG 30-JUL-2004  
DEFINITION Gorilla gorilla gorilla clone CH255-73E2, WORKING DRAFT SEQUENCE, 6  
ordered pieces.

AC149549

HTG: HTGS\_PHASE2; HTGS DRAFT.

Gorilla gorilla gorilla (lowland gorilla)

Gorilla gorilla gorilla

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.

1 (bases 1 to 202169)

## AUTHORS

Walker, M.L., Osoegawa, K., Friedman, C., de Jong, P.J., Trask, B.J.,  
Antonellis, A., Ayale, K., Benjamin, B., Blakesley, R.W.,  
Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,  
Coleman, H., Daki, N., Engle, J., Guan, X., Gupta, J., Haghighi, P.,  
Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J.R., Jones, C.,  
Karlings, E., Kim, H., Kwong, P., Laric, P., Larson, S., Lee-Lin, S.-O.,  
Legaspi, R., Madden, M., Maduro, O.L., Maduro, V.B., Margulies, E.H.,  
Mastello, C., Maskeri, B., McDowell, J., Mullikin, J.C., Paguirigan, C.,  
Park, M., Portnoy, W.E., Prasad, A., Puri, O., Reddix-Dugue, N.,  
Schandler, K., Schueler, M.G., Shah, K., Sison, C., Stantripop, S.,  
Thomas, J.W., Thomas, P.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D.,  
Young, A. and Green, E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 202169)

Direct Submission

Submitted (08-JUN-2004) NIH Intramural Sequencing Center, 8717

Grovmont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 202169)

Green, E.D.

Direct Submission

Submitted (30-JUL-2004) NIH Intramural Sequencing Center, 8717

Grovmont Circle, Gaithersburg, MD 20877, USA

On Jul 30, 2004 this sequence version replaced gi:48427674.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>Contact: [nisc.zoo@nhgri.nih.gov](mailto:nisc.zoo@nhgri.nih.gov)

----- Project Information

Center project name: eqn

Center clone name: 073E02

The sequence data in this record represents an 'enhanced'  
version of a Phase 2 submission. Specifically, the indicated  
order and orientation of each sequence contig has been  
established using one or more of the following: read-pair  
data from individual subclones, overlaps with neighboring  
clones, alignment with available reference sequence (e.g.,

human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 200572 bases at least Q40  
Consensus quality: 201163 bases at least Q30  
Consensus quality: 201502 bases at least Q20  
Insert size: 176000; agarose-ff  
Insert size: 201669; sum-of-contigs  
Quality coverage: 10.25x in Q20 bases; agarose-ff  
Quality coverage: 8.94x in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs \* are represented as runs of N. The order of the pieces \* is believed to be correct as given, however the sizes \* of the gaps between them are based on estimates that have \* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and \* the accession number will be preserved.  
\* 1 15312: contig of 15312 bp in length  
\* 15313 15412: gap of unknown length  
\* 15413 73350: contig of 57938 bp in length  
\* 73351 73450: gap of unknown length  
\* 73451 80732: contig of 7282 bp in length  
\* 80733 80832: gap of unknown length  
\* 80833 133153: contig of 52321 bp in length  
\* 133154 133253: gap of unknown length  
\* 133254 159871: contig of 26618 bp in length  
\* 159872 202169: gap of unknown length  
\* 159972 202169: contig of 42198 bp in length.

## FEATURES

source  
1..202169  
/organism="Gorilla gorilla gorilla"  
/mol\_type="genomic DNA"  
/sub\_species="gorilla"  
/db\_xref="taxon:9595"  
/clone\_lib="CH255-73E2"  
/clone\_lib="CH255"  
/note="BAC resource: <http://bacpac.chori.org/>"  
misc\_feature  
1..15312  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"  
15413..73350  
/note="assembly\_fragment"  
73451..80732  
/note="assembly\_fragment"  
80833..133153  
/note="assembly\_fragment"  
133254..159871  
/note="assembly\_fragment"  
159972..202169  
/note="assembly\_fragment  
clone\_end:T7  
vector\_side:right"

## ORIGIN

Query Match 8.0%; Score 160.4; DB 2; Length 202169;  
Best Local Similarity 49.6%; Pred. No. 5.2e-10;  
Matches 551; Conservative 0; Mismatches 546; Indels 13; Gaps 5;  
Qy 496 ATTGTTTATATCTTAAATTAATAATTTGTCCTCTTCAAAATTTTCATATAGTTATTTAT 555  
Db 11426 AATATTTATATATATTTCTGTTAATAATATATATATATATATATATATATATATAT 11367

Qy 556 ATATTACTTTTTCTCTATTTCTATTTCTATTTCTATTTCTCAAAATTTATTTATTTATGTCATATGTA 615  
Db 11366 TATATTTTATATAAATATTTTATATATATATTTTAAATTTTATATATTTTAAATTTTATATGCAATATA 11307  
Qy 616 AAGTACATTTATTTT---TTGCTATATATTTTAAATATTTCTAAATTTTCTAAATTTTAAAAAAGACT 672  
Db 11306 TATATAATATATATAAATATTTTATATGAATATATATATAAATATATATAAATATTTTAAATATTT 11247  
Qy 673 GATATGAAAAATTTATTTCTTTTAAAGCTATATCATTTTATATATATATATATATATTTTCTTTTCTT 732  
Db 11246 TATATATGAATATATATATAAATATATATAAATATTTTATATAAATATATATAAATATA 11187  
Qy 733 TTTCTTTCTAT-TTTCTATTTCAATTTTAAAGAAATTTTGTAAATTTTGTAAATTTTATTTATCAA 791  
Db 11186 TATATAAATATTTATATATGAATATATATAAATATATATAAATATTTTATATAAATA 11127  
Qy 792 TTTTAAAAATATTTTACTTTTATATGTTTCTTTTCAATTTTGTAAACAAATCATATCAT 851  
Db 11126 TATATAAATATTTTATATAAATATATATAAATATAGATATATAAATATATATATATCTATTT 11067  
Qy 852 TATGATTTGAAAGAGAGAGAAATTTGACAGTGAGTAAATAGTGAGAGAAAAAATGTTTAT 911  
Db 11066 TATATAAATATATAAATATATATATCTATTTTATATAAATATATAAATATAAAT 11007  
Qy 912 TTTCTAAAAAACCTTAAACAAACATGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 971  
Db 11006 ATATTAATAAATAAATATAAATATATATATAAATATAAATATAAATATAAATATA 10947  
Qy 972 ATTTTCTCTTTATCT 1031  
Db 10946 ATAAATATATATTTTATATATTTTGTATATATTTATATAAATATATAAATACATATTTATA 10887  
Qy 1032 ACTCTCTTTATTTTCT 1091  
Db 10886 TATATATTTATATTTTATATTTTATATTTTATATTTTATATTTTATATATTTTATATA 10827  
Qy 1092 TTGGGGCATGCTTTATCATCT 1151  
Db 10826 TATTATATTTTATATATTTTGTATATATATATAAATATATATATATATATATATATATA 10767  
Qy 1152 AAGCATTTGGCAAGTCT 1211  
Db 10766 TAATTATA-----TATTATATATATATATATTTTATATATTTTATATAAATATAAATA 10714  
Qy 1212 TGATTTCAGTATATCTATTTGCAATGTGTAAGTCTTTCCACAAATACATAAATTTAAATTTAA 1271  
Db 10713 TATATAAATATGTAACATATATTTTATATAAATATAAATATAAATATGTAACATATGTAATTTA 10654  
Qy 1272 TCTTAAATAAATAAAGGATAAATAATTTTCTTTCTCTCAAAAATTTAAATATATGTTATTT 1331  
Db 10653 TATATAAATATATAAATATGTAACATATATATTTTATATAAATATATAAATATGTAACATA 10594  
Qy 1332 TTTGTTTATAGTATATTTTCGAATAAATCTAAATATATATATATATGATATGATTTTATATTTGAT 1391  
Db 10593 TATATTTATATAAATATATAAATATGTAACATATATTTTATATAAATATATAAATA 10534  
Qy 1392 AAACATATAATCAATATTTAAATATGATATTTTATATATAGTGTGTCACACATAAATTTTAT 1451  
Db 10533 TGTAACATATATTTTATGTAATAAATATAAATATAAATATAAATAAATAAATAAATAAATA 10475  
Qy 1452 AAGGATAAATAATGATAAATAAATTTTAAATATTTTATATATTTTATATTTTACGAGAAAAAAA 1511  
Db 10474 AAATAT-ATRAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATAAATA 10416  
Qy 1512 ATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACTTATGATTAATTCATAAATTC 1571  
Db 10415 TAAATATAAATAAATAAATAAATATACATTTGTTATATTTTATATATATCTATATAAATATA 10356  
Qy 1572 TATATGATATTTTGAATTTAAAAACAGATA 1601  
Db 10355 TATTACATATCTATATATAAATAAATA 10326





```
Db 26423 ATTTAAATATATATATTTATTTATATATATTTATTAATTAATTAATTAATTTATTTATTTATTTATTT 26482
Qy 1407 ATTAATATATGATATTTTATATATAGTTGTACACATAATTTATAGGATAAAAAATAT 1466
Db 26483 ATAAATATATATATTTTATTTATTTATTAATATTTAAATTTGATTAATATTTTCTCATTTAT 26542
Qy 1467 GATAAAATAAATTTTAAATATTTTATTTATTTTACGAGAAAAAAAT--ATTTTAGCCA 1523
Db 26543 TATTTATATTAATATTCATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 26602
Qy 1524 TAAATAATGACAGCATATTTTACAACCTTGTAGTAATTCATAAAATTCCTATATGATATTT 1583
Db 26603 TATTTAAATTTATTTATTTTATTTTATTTTATTTTAAATTAATTAATTAATTTTATTAATA 26662
Qy 1584 TGAATTTAAACACAGATAATCGTTAAGGAGGAATCCTAGTCATCTCTGCAATTTGT 1643
Db 26663 TATATTTTAAATTTTAAATATATTTTAAATAAATTTTATAATGATTTAAATTTGATATATTA 26722
Qy 1644 TTTTCATGCAACAGAAA 1661
Db 26723 TATAATATATATATATA 26740
```

```
RESULT 10
CR626872/c
LOCUS CR626872 180096 bp DNA linear HTG 27-JUL-2004
DEFINITION Danio rerio clone DKEY-38L22, *** SEQUENCING IN PROGRESS ***, 12
unordered pieces.
ACCESSION CR626872
VERSION CR626872.1 GI:50724874
KEYWORDS HTG; HTGS PHASE1.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 180096)
Sims.S.
Direct Submission
Submitted (23-JUL-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk
----- Project Information
Center project name: zk38L22
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 175236 bases at least Q40
Consensus quality: 176399 bases at least Q30
Consensus quality: 177283 bases at least Q20
Insert size: 178996; sum-of-contigs
Insert size: 184477; 3.4% error; agarose-fp
Quality coverage: 4.66x in Q20 bases; sum-of-contigs Quality
coverage: 4.66x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 19704: contig of 19704 bp in length
* 19705 19804: gap of 100 bp
* 19805 31589: contig of 11785 bp in length
* 31590 31689: gap of 100 bp
```

```
* 31690 38161: contig of 6472 bp in length
* 38162 38261: gap of 100 bp
* 38262 60431: contig of 22170 bp in length
* 60432 60531: gap of 100 bp
* 60532 85983: contig of 25452 bp in length
* 85984 86083: gap of 100 bp
* 86084 91753: contig of 5670 bp in length
* 91754 91853: gap of 100 bp
* 91854 114485: contig of 22632 bp in length
* 114486 125670: contig of 11085 bp in length
* 125671 125770: gap of 100 bp
* 125771 132626: contig of 6856 bp in length
* 132627 132726: gap of 100 bp
* 132727 142478: contig of 9652 bp in length
* 142479 142479: gap of 100 bp
* 142479 175518: contig of 33040 bp in length
* 175519 175618: gap of 100 bp
* 175619 180096: contig of 4478 bp in length.

FEATURES
             Location/Qualifiers
     source          1..180096
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-38L22"
                     /clone_lib="DanioKey"
     misc_feature    1..19704
                     /note="assembly_fragment:00533
                     fragment_chain:1"
     misc_feature    19805..31589
                     /note="assembly_fragment:00302
                     fragment_chain:1"
     misc_feature    31690..38161
                     /note="assembly_fragment:00091
                     fragment_chain:1"
     misc_feature    38262..60431
                     /note="assembly_fragment:00381
                     fragment_chain:1"
     misc_feature    60532..85983
                     /note="assembly_fragment:00844
                     fragment_chain:1"
     misc_feature    86084..91753
                     /note="assembly_fragment:00055
                     fragment_chain:1"
     misc_feature    91854..114485
                     /note="assembly_fragment:00687
                     fragment_chain:1"
     misc_feature    114586..125670
                     /note="assembly_fragment:00234
                     fragment_chain:1"
     misc_feature    125771..132626
                     /note="assembly_fragment:00133
                     fragment_chain:1"
     misc_feature    132727..142378
                     /note="assembly_fragment:00176
                     fragment_chain:1"
     misc_feature    142479..175518
                     /note="assembly_fragment:01036
                     fragment_chain:1"
     misc_feature    175619..180096
                     /note="assembly_fragment:00023.0"

ORIGIN
Query Match          7.9%; Score 159.8; DB 2; Length 180096;
Best Local Similarity 43.3%; Pred.No. 6.3e-10;
Matches 479; Conservative 0; Mismatches 624; Indels 3; Gaps 3;
QY 498 TGTATTATCTTAAATAAAATATTGTCCTTCAAATTTTCATATAGTTAATTATAT 557
      |||||
Db 115252 TATATATATATATATATATATATATATATATATATATATATATATATATATA 115193
QY 558 ATTACTTTTTCCTATCTTATCTTATTTTCAAATTTATTTATGTCATATGATAA 617
      |||||
```









REFERENCE TITLE JOURNAL AUTHORS	Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emrling,S., Escotto,M., Falle,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsif,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,I.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisleg,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguy,M., Okwunu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shoostari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
REFERENCE TITLE JOURNAL AUTHORS	2 (bases 1 to 143331) Worley,K.C. Direct Submission Submitted (04-APR-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE TITLE JOURNAL AUTHORS	3 (bases 1 to 143331) Worley,K.C. Direct Submission Submitted (27-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE TITLE JOURNAL AUTHORS	4 (bases 1 to 143331) Worley,K.C. Direct Submission Submitted (28-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE TITLE JOURNAL AUTHORS	5 (bases 1 to 143331) Worley,K.C. Direct Submission Submitted (12-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE TITLE JOURNAL AUTHORS	6 (bases 1 to 143331) Worley,K.C. Direct Submission Submitted (27-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE TITLE JOURNAL AUTHORS	7 (bases 1 to 143331) Worley,K.C. Direct Submission Submitted (31-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE TITLE JOURNAL AUTHORS	8 (bases 1 to 143331) Worley,K.C. Direct Submission Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
COMMENT	of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA On Jul 17, 2001 this sequence version replaced gi:14578078. INFORMATION: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a> or email <a href="mailto:gc-help@bcm.tmc.edu">gc-help@bcm.tmc.edu</a>  CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.  ANNOTATION OF FEATURES: STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts. Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.  SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.  QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <a href="http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html">http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html</a> .  QUALSTAT-REPORT. Location/Qualifiers 1. .143331 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="12" /clone="RP11-13F9" /complement(1. .2005) /note="overlaps bases 1. .2005 of clone AC079409" /function="clone overlap" 511. .703 /rpt_family="LTR16A" complement(825. .1128) /rpt_family="AluY" complement(3191. .3292) /rpt_family="L2" 3856. .3887 /rpt_family="AT-rich" complement(4061. .4369) /rpt_family="LTR33A" 4457. .5350 /rpt_family="LIM1" 5351. .5373 /rpt_family="T-rich" complement(5374. .6088) /rpt_family="LIP7" 6109. .6146 /rpt_family="(TTTTA)n" 7538. .7716 /standard_name="61535" 12648. .13327 /rpt_family="LIM1"

```
repeat_region 13353..13387
/rpt_family="AT_rich"
repeat_region 13423..15610
/rpt_family="LIM1"
repeat_region 15611..15710
/rpt_family="FLAM_C"
repeat_region 15718..15889
/rpt_family="TA)n"
repeat_region 15909..16212
/rpt_family="TA)n"
repeat_region complement(16213..16263)
/rpt_family="FLAM"
misc_feature 16261..17578
/function="unresolved tandem repeat"
repeat_region 16264..16399
/rpt_family="TA)n"
repeat_region 16405..16446
/rpt_family="AT_rich"
repeat_region 16447..17315
/rpt_family="TA)n"
repeat_region 17336..17504
/rpt_family="TA)n"
repeat_region 17599..17636
/rpt_family="TA)n"
repeat_region 17648..17764
/rpt_family="AluJo/FRAM"
repeat_region 17774..17807
/rpt_family="AT_rich"
repeat_region 17811..18491
/rpt_family="L1"
repeat_region 18494..18612
/rpt_family="AluJb"
repeat_region 18613..19807

Query Match
Best Local Similarity 7.9%; Score 158.4; DB 9; Length 143331;
Matches 543; Conservative 0; Mismatches 531; Indels 16; Gaps 5;

Qy 480 TTACATGAAAGAAATATGTTATATCTCTTAATTAATAAATATGTCCTCTCAAAATTTT 539
Db 16468 TTATTTATATTTATATATATATATATATATATATATATATATATATATATATAT 16527

Qy 540 CATATAGTAATATATATATATATATATATATATATATATATATATATATATATAT 599
Db 16528 TATTTATATATACATATATATATATATATATATATATATATATATATATATATAT 16587

Qy 600 TATTTATGATGTAAGTACATATATATATATATATATATATATATATATATATATAT 659
Db 16588 TATATATATTTTATATATATATATATATATATATATATATATATATATATATAT 16647

Qy 660 TTAAGAAAGACTGATATGAAAAA-----TTTATCTTTTAAAGCTATATCATTTTAT 715
Db 16648 TTATGATATATTTATATATATATATATATATATATATATATATATATATATATAT 16707

Qy 716 ATACTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 775
Db 16708 ATATATATTTTATATATATATATATATATATATATATATATATATATATATAT 16767

Qy 776 AATTTTATTTATCAATTTATATAAATAATTTTATCTTTATATATGTTTTCACATTTTGT 835
Db 16768 TATGTAATTTTATATATATATATATATATATATATATATATATATATATATATAT 16827

Qy 836 AAACAAATCATATCATATATGATGAAAGAGAGGAAATTCACAGTCAAGTAAGTATGA 895
Db 16828 TTATATAT-ATTTTATATATATTTATATATATATATATATATATATATATATATAT 16886

Qy 896 GAAAAAATGTTATTTCTTAAAAAACCCTAAACAATGATCTACTCTCTATTTTC 955
Db 16887 TTATATATATATATATATTTATACATATTTTATACATATTCATATATATATATAT 16946

Qy 956 ATCTATCTCTCATTTTCATTTT---TCTCTTTATCTCTTTCTTTCTTTCTTTCTTTCT 1012
Db 16947 TATTTTATATATATATATATATATATATATATATATATATATATATATATATAT 17006
```

```
Qy 1013 TTTCACTAATTAATATATTTTACTCTCTTTATTTTCTCTCTATCCCTCTCTTATTTCCAC 1072
Db 17007 ATTTATATATATATATTTTATATATATTTTATATATTTTATATATATATATATATTTG 17066

Qy 1073 TCATATATACACTCCAAAATTTGGGCGCATGCTTTTACTACTACTCTATCTCTCCACTAA 1132
Db 17067 ATATATATTTATATATTTATATATATTTTATATATATTTATATATATATATATATTTCA 17126

Qy 1133 TCATTTAAATGAAGACTGAAAGACATTTGGCAAGTCTCTCCCTCCCTCTCAAGTATTT- 1187
Db 17127 TTTTATATATTTATATATATATATATATTTTATATATATTTTATATATATTTATATA 17186

Qy 1188 --TCCAACTCAGCATTTGGCATCTCGATTCAGTATATCTATTCGATGCTGTAAAGTCT 1245
Db 17187 TATTTATATATATTTTATATATATTTTATATATATTTTATATATATTTATATATATTT 17246

Qy 1246 TTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAATAATTTTTTTTT 1305
Db 17247 ATTTATATATATTTTATATATATATATATTTTATATATATTTTATATATATTTATATA 17306

Qy 1306 CTTCATAAAATTTAAATATATGT-TATTTTTTCTTTAGATGATATTTCCGAATAAATCTAA 1364
Db 17307 TTATATATAGTTTATATATTTATATATTTTATATATATTTTATATATATTTTATATA 17366

Qy 1365 ATATGATAATGATTTTTTATATTTGATTTAAACATATATAATCAATATATAATATGATAT 1424
Db 17367 ATATTTTATATATATTTTATATTTTATATATATTTTATATATTTTATATATATATTT 17426

Qy 1425 TTATATAGTTGTACACATAAATTTTATAAGGATAAATAAATATGATAAATAAATTTTAA 1484
Db 17427 TTATATATTTTATATATATATATTTTATATATTTTATATATAAATAAATAATATAT 17486

Qy 1485 ATATTTTATATATTTTACGAGAAAAAATAATTTTATAGCCATAAATAAATGCACGATAT 1544
Db 17487 TATTTATATATATATATATATATATATATATATATATTTTTCNGTAGATACATATA 17546

Qy 1545 TTACAACCTT 1554
Db 17547 ATATACATTT 17556
```

RESULT 14  
CR759967/c  
LOCUS

DEFINITION CR759967 226108 bp DNA linear HTG 11-SEP-2004  
Danio rerio clone DKEY-94P13, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 6  
unordered pieces.

ACCESSION CR759967  
VERSION CR759967.2 GI:51988233  
KEYWORDS HTG; HTGS PHASE1.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
REFERENCE 1 (bases 1 to 226108)  
Sims,S.  
Direct Submission

JOURNAL Submitted (10-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Sep 11, 2004 this sequence version replaced gi:51965792.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC

Web site: <http://www.sanger.ac.uk>  
Contact: zfish-help@sanger.ac.uk  
----- Project Information  
Center project name: zK94P13  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 223262 bases at least Q40

Consensus quality: 223728 bases at least Q30									
Consensus quality: 224330 bases at least Q20									
Insert size: 225608; sum-of-contigs									
Insert size: 219705; 1.4% error; agarose-fp									
Quality coverage: 11.47x in Q20 bases; sum-of-contigs Quality coverage: 11.83x in Q20 bases; agarose-fp									
-----									
* NOTE: This is a 'working draft' sequence. It currently									
* consists of 6 contigs. The true order of the pieces									
* is not known and their order in this sequence record is									
* arbitrary. Gaps between the contigs are represented as									
* runs of N, but the exact sizes of the gaps are unknown.									
* This record will be updated with the finished sequence									
* as soon as it is available and the accession number will									
* be preserved.									
-----									
1 5476: contig of 5476 bp in length									
* 5477 5576: gap of 100 bp									
* 5577 34821: contig of 29245 bp in length									
* 34822 34921: gap of 100 bp									
* 34922 59876: contig of 24955 bp in length									
* 59877 59976: gap of 100 bp									
* 59977 138375: contig of 78399 bp in length									
* 138376 138475: gap of 100 bp									
* 138476 176066: contig of 37591 bp in length									
* 176067 176166: gap of 100 bp									
* 176167 226108: contig of 49942 bp in length.									
FEATURES									
source									
1..226108									
/organism="Danio rerio"									
/mol_type="genomic DNA"									
/db_xref="taxon:7955"									
/clone="DKEY-94P13"									
/clone_lib="DanioKey"									
1..5476									
/note="assembly_fragment:00022									
fragment_chain:1									
5577..34821									
/notes="assembly_fragment:00494									
fragment_chain:1									
34922..59876									
/note="assembly_fragment:00083									
fragment_chain:2									
59977..138375									
/notes="assembly_fragment:02246									
fragment_chain:2									
138476..176066									
/note="assembly_fragment:00979"									
176167..226108									
/notes="assembly_fragment:01544.0"									
ORIGIN									
Query Match 7.9%; Score 158.4; DB 2; Length 226108;									
Best Local Similarity 43.3%; Pred. No. 8.4e-10;									
Matches 479; Conservative 0; Mismatches 623; Indels 4; Gaps 4;									
Qy	498	TGTTTATATTCTTAAATTAATAATATTGTCCTCTTAAATTTTCATATAGTTAATTATTAT	557						
Db	60508	TA	60449						
Qy	558	ATTACTTTTTCTTATCTTATAGTTCTATTCTTTTCAAAATTAATTTATGCAATATGTA	617						
Db	60448	TA	60389						
Qy	618	GTACATATATTTTTTGCTATATACCTTAATTAATTTCTAAATTTATTAAGAAAGACTG	677						
Db	60388	TA	60330						
Qy	678	GAAAAATTTATTCCTTTTAAAGCTATCATCTTTTATATATACCTTTTCTTTCTTTCTT	737						
Db	60329	ATA	60270						
Qy	738	TCATTTTCTATTCAATTTTAATAAGAAATAAAT-TTTGTAAAAATTTTATTTATCAATTA	796						

JOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 129240)  
Biren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Barna, N., Bastien, V., Bada, F., Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., O'Neil, D., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strause, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Direct Submission  
Submitted (13-OCT-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 129240)  
Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Direct Submission  
Submitted (02-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 129240)  
Biren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL

Direct Submission  
Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On Jul 30, 2002 this sequence version replaced gi:21306731.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WtBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L11383  
Center clone name: 3118\_D\_11  
-----

FEATURES  
source

Location/Qualifiers  
1. 129240  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="CTD-3118D11"  
/clone\_lib="CITD2 Human BAC"  
1. 224  
/rpt\_family="AluSg/x"  
225. 403  
/rpt\_family="AluJb"  
669. 855  
/rpt\_family="L1MC3"  
1519. 1600  
/rpt\_family="AT rich"  
complement(1619..1895)  
/rpt\_family="AluSc"  
complement(1897..2197)  
/rpt\_family="AluSg"  
2310. 2593  
/rpt\_family="Charlie2"  
complement(2598..3451)  
/rpt\_family="L1MA4A"  
2966. 2969  
/notes="single clone coverage"  
3483. 3713  
/rpt\_family="L1MDa"  
complement(3714..4017)  
/rpt\_family="AluSg"  
3739. 3748  
/notes="<30 qual SNGL region"  
3785. 3811  
/note="<30 qual SNGL region"  
4018. 4255  
/rpt\_family="L1MDa"  
4267. 4982  
/rpt\_family="L1MDa"  
4974. 5464  
/rpt\_family="L1MD2"  
5827. 6015  
/rpt\_family="MIR"  
8136. 8345  
/rpt\_family="MIR"  
8552. 8737  
/rpt\_family="L2"  
9347. 9644  
/rpt\_family="AluSx"  
complement(10188..10324)  
/rpt\_family="MIR3"  
10770. 11002  
/rpt\_family="MIR"  
complement(11116..11638)  
/rpt\_family="MER1A"  
12153. 12222  
/rpt\_family="L2"  
complement(12668..12878)  
/rpt\_family="MIR"  
complement(13676..14221)





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 21:53:41 ; Search time 781.833 Seconds  
(without alignments)  
15234.088 Million cell updates/sec

Title: US-10-776-889-1

Perfect score: 2012

Sequence: 1 atcttagcccttgattata.....aactaagaagcaatacttca 2012

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2012	100.0	2012	13	ADR16015 Soybean a
2	2012	100.0	2012	13	ADR20130 Soybean a
3	1883	93.6	1883	13	ADR16027 Soybean a
4	1719	85.4	1719	13	ADR16028 Soybean a
5	1553	77.2	1553	13	ADR16029 Soybean a
6	1367	67.9	1367	13	ADR16030 Soybean a
7	1160	57.7	1160	13	ADR16031 Soybean a
8	967	48.1	967	13	ADR16032 Soybean a
9	770	38.3	770	13	ADR16033 Soybean a
10	584	29.0	584	13	ADR16034 Soybean a
11	425	21.1	425	13	ADR16035 Soybean a
12	174	8.6	174	13	ADR16036 Soybean a
13	156	7.8	1548	6	ABL34155 Human imm
14	150.2	7.5	8056	8	ABZ10246 Haematopo
15	148.6	7.4	8056	8	ABZ10246 Haematopo
16	147.4	7.3	7597	6	ABL33013 Human imm
17	142.2	7.1	8056	8	ABZ10100 Haematopo
18	136.8	6.8	5286	13	ADS89552 Oligonuc
19	136	6.8	115218	8	ACA64845 Human HNR
20	135.8	6.7	110000	13	ABD32968_6 Continuation (7 of

c	21	135.2	6.7	5286	13	ADS89278	Ad889278 Oligonuc
	22	133.8	6.7	5286	13	ADS89278	Ad889278 Oligonuc
	23	133.8	6.7	5286	13	ADS89552	Ad889552 Oligonuc
	24	133.6	6.6	8056	8	ABZ10100	Abz10100 Haematopo
c	25	133	6.6	158001	12	ADL17884	Adl17884 Human pho
	26	131.8	6.6	7814	4	AA546530	Aas46530 Tumour su
	27	128.4	6.4	7498	6	ABL32257	Abi32257 Human imm
c	28	127.8	6.4	99764	13	ABD32992	Abd32992 Human can
	29	124.4	6.2	6175	6	ABL33307	Abi33307 Human imm
	30	123.8	6.2	6419	6	ABL32267	Abi32267 Human imm
	31	123.8	6.2	12237	6	ABL34358	Abi34358 Human imm
c	32	122.4	6.1	1300	12	ADP85917	Adp85917 Synthetic
	33	122.4	6.1	110000	13	ABD32968_6	Continuation (7 of
c	34	120.6	6.0	1000	12	ADQ62833	Adq62833 Homopoly-
	35	120.6	6.0	1000	12	ADQ62832	Adq62832 Homopoly-
	36	120	6.0	6109	6	ABL32326	Abi32326 Human imm
	37	120	6.0	6109	6	AA561077	Aas61077 Human gen
	38	119	5.9	34548	6	ABL70604	Abi70604 Chemical
	39	118.2	5.9	7442	4	AA546686	Aas46686 Tumour su
c	40	118	5.9	115218	8	ACA64845	Ac64845 Human HNR
	41	116.4	5.8	11222	10	ADBS4318	Adbs4318 Pretreat
c	42	116.4	5.8	110000	10	ADH10017_0	Adh10017 Human chr
	43	116.4	5.8	158001	12	ADL17884	Adl17884 Human pho
	44	116	5.8	8222	8	ACF62816	Acf62816 Colon can
	45	116	5.8	8222	8	ACF62794	Acf62794 Colon can

ALIGNMENTS

RESULT 1

ADR16015  
ID ADR16015 standard; DNA; 2012 BP.

XX ADR16015;

AC ADR16015;

XX DT 18-NOV-2004 (first entry)

DE XX Soybean annexin promoter.

XX KW Soybean; annexin promoter; ds; dicotyledonous plant; plant.

XX OS Glycine max.

XX PN US2004158052-A1.

XX PD 12-AUG-2004.

XX PF 11-FEB-2004; 2004US-00776889.

XX PR 12-FEB-2003; 2003US-0446833P.

XX PA (KINN/) KINNEY A J.

XX PI Kinney AJ, Liu Z;

CC soybean. The method of regulating expression of at least one heterologous  
CC nucleotide sequence in a plant comprises transforming a plant cell with  
CC the recombinant expression construct, growing fertile mature plants from  
CC a transformed plant cell and selecting plants comprising a transformed  
CC plant cell expressing the heterologous nucleotide sequence.  
CC Alternatively, plants comprising a transformed plant cell expressing the  
CC heterologous nucleotide sequence during early seed development are  
CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin  
CC and P34 promoters. The promoters are useful in regulating or promoting  
CC expression of at least one heterologous nucleic acid fragment in plants.  
XX This sequence represents the soybean annexin promoter of the invention.  
SQ Sequence 2012 BP; 695 A; 326 C; 213 G; 778 T; 0 U; 0 Other;

Query Match		100.0%;	Score 2012;	DB 13;	Length 2012;
Best Local Similarity		100.0%;	Pred. No. 3.5e-225;		
Matches 2012;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATCTTAGGCCCTTGATTATATGTTGGTTTGTAGATGGATTACATGCAAGTTTTTATTTCAAT	60		
Db	1	ATCTTAGGCCCTTGATTATATGTTGGTTTGTAGATGGATTACATGCAAGTTTTTATTTCAAT	60		
Qy	61	CCCTTTTCCCTTGGTAATACTGACCAAGAACCAAGAAAAAAGAAAGGATCA	120		
Db	61	CCCTTTTCCCTTGGTAATACTGACCAAGAACCAAGAAAAAAGAAAGGATCA	120		
Qy	121	TTTTGAAGGATATTTTCCGCTCTATTCAAATCTGATTTTACCAAAAAAAGCTGAT	180		
Db	121	TTTTGAAGGATATTTTCCGCTCTATTCAAATCTGATTTTACCAAAAAAAGCTGAT	180		
Qy	181	TTTTCTACACTCTCAAGCTTTGTTTTGCTTCGACTCTCATGATTTCTTCATATGCC	240		
Db	181	TTTTCTACACTCTCAAGCTTTGTTTTGCTTCGACTCTCATGATTTCTTCATATGCC	240		
Qy	241	AATCACTCTATTATAAATGSCATAAGGTAGTGGAACAATTGCAAAAGCTTGTCATCAAA	300		
Db	241	AATCACTCTATTATAAATGSCATAAGGTAGTGGAACAATTGCAAAAGCTTGTCATCAAA	300		
Qy	301	AGCTTGAAGTCAACAATTAATGTTTTTTCATGCTTTCAAAATTTATCTGCACCCCTPAGC	360		
Db	301	AGCTTGAAGTCAACAATTAATGTTTTTTCATGCTTTCAAAATTTATCTGCACCCCTPAGC	360		
Qy	361	TATTAATCTAACACTCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCAATGTCATTA	420		
Db	361	TATTAATCTAACACTCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCAATGTCATTA	420		
Qy	421	ATTTCCCGTGACTATTTTGGCTTTGACTCCAACTGGCCCGGTACATCCGTCCTCAT	480		
Db	421	ATTTCCCGTGACTATTTTGGCTTTGACTCCAACTGGCCCGGTACATCCGTCCTCAT	480		
Qy	481	TACATGAAAGAAATATTGTTTATATTCTTAATTAATAAATATTGTCCTCTTCAAAATTTTC	540		
Db	481	TACATGAAAGAAATATTGTTTATATTCTTAATTAATAAATATTGTCCTCTTCAAAATTTTC	540		
Qy	541	ATATAGTTAAATTAATTAATTAATTTTCTCTATTCTATTAGTTCTATTTTCAAAATATT	600		
Db	541	ATATAGTTAAATTAATTAATTAATTTTCTCTATTCTATTAGTTCTATTTTCAAAATATT	600		
Qy	601	ATTTATGATGATGAAAGTACATTAATTTTGGCTATATATCTTAAATATTCTTAAATATT	660		
Db	601	ATTTATGATGATGAAAGTACATTAATTTTGGCTATATATCTTAAATATTCTTAAATATT	660		
Qy	661	TAAAAAAGACTGATATGAAAAATTTTATCTTTTAAAGCTATATCATTTTATATATACT	720		
Db	661	TAAAAAAGACTGATATGAAAAATTTTATCTTTTAAAGCTATATCATTTTATATATACT	720		
Qy	721	TTTTCTTTTCTTTTCTTTTCTATTCTATTCTATTCTATTCTATTCTATTCTATTCT	780		
Db	721	TTTTCTTTTCTTTTCTTTTCTATTCTATTCTATTCTATTCTATTCTATTCTATTCT	780		
Qy	781	TTATTTATCAATTTATAAAAATTTTACTTTATATGTTTTTTCACATTTTGTGTAACA	840		
Db	781	TTATTTATCAATTTATAAAAATTTTACTTTATATGTTTTTTCACATTTTGTGTAACA	840		

Qy	841	AATCATATCATTTATGATTGAAGAGAGGAAATTTGACAGTGAGTAATAAGTGATGAGAAAA	900		
Db	841	AATCATATCATTTATGATTGAAGAGAGGAAATTTGACAGTGAGTAATAAGTGATGAGAAAA	900		
Qy	901	AAATGTTGTTATTTCTTAAAAAAAACCTAAACAAACATGTATCTCTCTATTTCATCTA	960		
Db	901	AAATGTTGTTATTTCTTAAAAAAAACCTAAACAAACATGTATCTCTCTATTTCATCTA	960		
Qy	961	TCTCTCATTTTCATTTTCTTTTATCTCTTTTCTTTTCTTTTATCATATCATTTTCACAT	1020		
Db	961	TCTCTCATTTTCATTTTCTTTTATCTCTTTTCTTTTCTTTTATCATATCATTTTCACAT	1020		
Qy	1021	TAATTAATTTTACTCTCTTTTATTTTCTCTCTATCCCTCTCTTATTTTCCACTCATATAT	1080		
Db	1021	TAATTAATTTTACTCTCTTTTATTTTCTCTCTATCCCTCTCTTATTTTCCACTCATATAT	1080		
Qy	1081	ACACTCCAAAAATTTGGGGCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1140		
Db	1081	ACACTCCAAAAATTTGGGGCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1140		
Qy	1141	ATGAAACTGAAAGAGCATTTGGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1200		
Db	1141	ATGAAACTGAAAGAGCATTTGGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1200		
Qy	1201	TGGCATCTGATTGATTTCAGTATATCTATTGTCATGTGTAAAGTCTTTTCCACAATACATAA	1260		
Db	1201	TGGCATCTGATTGATTTCAGTATATCTATTGTCATGTGTAAAGTCTTTTCCACAATACATAA	1260		
Qy	1261	CTATTAATTAATCTTAAATAAATAAAGGATAAATAATTTTTTTTTCTTCATAAAAAATAAA	1320		
Db	1261	CTATTAATTAATCTTAAATAAATAAAGGATAAATAATTTTTTTTTCTTCATAAAAAATAAA	1320		
Qy	1321	ATATGTTATTTTGTGTTAGATGTATATTCGAATAATCTTAAATATATGATATATGATTTT	1380		
Db	1321	ATATGTTATTTTGTGTTAGATGTATATTCGAATAATCTTAAATATATGATATATGATTTT	1380		
Qy	1381	TTATATTTGATTAACATATAATCAATATTAAATATGATATTTTTTTTATATATAGTGTGACA	1440		
Db	1381	TTATATTTGATTAACATATAATCAATATTAAATATGATATTTTTTTTATATATAGTGTGACA	1440		
Qy	1441	CATAATTTTATAGGATAAATAAATAATGATAAATAAATAAATAAATAAATAAATAAATAA	1500		
Db	1441	CATAATTTTATAGGATAAATAAATAATGATAAATAAATAAATAAATAAATAAATAAATAA	1500		
Qy	1501	GAGAAAAAATAATTTTATAGCCATAATAAATGACCAAGCATATTTTACAACTTTAGTAAT	1560		
Db	1501	GAGAAAAAATAATTTTATAGCCATAATAAATGACCAAGCATATTTTACAACTTTAGTAAT	1560		
Qy	1561	TCATAAATTCCTATATATGTTGAAATTTAAAAACAGATAATCGTTAAGGGAGGAATC	1620		
Db	1561	TCATAAATTCCTATATATGTTGAAATTTAAAAACAGATAATCGTTAAGGGAGGAATC	1620		
Qy	1621	CTACGTCTCTCTGTCATTTGTTTTCATGCAAAACAGAAAGGAGCAAAACACCTCA	1680		
Db	1621	CTACGTCTCTCTGTCATTTGTTTTCATGCAAAACAGAAAGGAGCAAAACACCTCA	1680		
Qy	1681	CCATGAATCACTCTTCCACACCATTTTTTCTAGCAAAACAGTCTCAACACTGAAGCCAGC	1740		
Db	1681	CCATGAATCACTCTTCCACACCATTTTTTCTAGCAAAACAGTCTCAACACTGAAGCCAGC	1740		
Qy	1741	TCCTCTTTCCTGTTTCTTTTCAACACTTTTCTTTTGAATAGTAGTATTTTTTTTTCACATG	1800		
Db	1741	TCCTCTTTCCTGTTTCTTTTCAACACTTTTCTTTTGAATAGTAGTATTTTTTTTTCACATG	1800		
Qy	1801	ATTTATTAACGTGCCAAAAGATGCTTATTTGAATAGAGTGACATTTTGTATGTACTACTA	1860		
Db	1801	ATTTATTAACGTGCCAAAAGATGCTTATTTGAATAGAGTGACATTTTGTATGTACTACTA	1860		
Qy	1861	ATTAGAACATGAAAAAGCATTTGTTTAAACACGATAATCTCTGTGAAGGCGTTAACTCCAAA	1920		
Db	1861	ATTAGAACATGAAAAAGCATTTGTTTAAACACGATAATCTCTGTGAAGGCGTTAACTCCAAA	1920		

QY 1921 GATCAATTTCACTATATAAATTTGACGAAGCAAAATGAATTCACATAGCTGAGAG 1980  
Db GATCAATTTCACTATATAAATTTGACGAAGCAAAATGAATTCACATAGCTGAGAG 1980  
QY 1981 AAGGAAAGGTTAACTAAGAACAATCTTCA 2012  
Db AAGGAAAGGTTAACTAAGAACAATCTTCA 2012

RESULT 2

ADR20130  
ID ADR20130 standard; DNA; 2012 BP.  
XX ADR20130;  
AC ADR20130;  
DT 18-NOV-2004 (first entry)  
XX Soybean annexin promoter nucleotide sequence SEQ ID NO:3.  
KW oilseed plant; mature seed; seed fatty acid profile;  
KW polyunsaturated fatty acid; oil; food; food product; beverage;  
KW infant formula; nutritional supplement; pet food; animal feed;  
KW whole bean soy product; aquaculture food product; soybean; annexin;  
KW promoter; ds.  
XX Glycine max.  
OS WO2004071467-A2.  
PN 26-AUG-2004.  
XX 12-FEB-2004; 2004WO-US005758.  
XX 12-FEB-2003; 2003US-0446941P.  
PR (DUPO ) DU PONT DE NEMOURS & CO E I.  
PA Kinney AJ, Cahoon EB, Damude HG, Hitz WD, Kolar CW, Liu Z;  
PI WPI; 2004-625770/60.  
PT Novel oilseed plant useful for mature seeds in which total seed fatty  
PT acid profile comprises high polyunsaturated fatty acid.  
XX

Example 1; SEQ ID NO 3; 132pp; English.

The present invention describes an oilseed plant (I) that produces mature seeds in which the total seed fatty acid profile comprises at least 1.0%, 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 60% or more of at least one polyunsaturated fatty acid having 20 or more carbon atoms and 5 or more carbon-carbon double bonds, and ratio of docosahexanoic acid:icosapentaenoic acid (DHA:EPA) is 1:100-860:100 or 1:100-110:100, where total seed fatty acid profile further comprises less than 2% arachidonic acid. Also described: (1) seeds (II) obtained from (I); (2) oil (III) obtained from (II); (3) a recombinant construct (IV) for altering the total fatty acid profile of mature seeds of an oilseed plant, comprising two or more promoters, where each promoter is operably linked to a nucleic acid sequence encoding a polypeptide (e.g., delta-4 desaturase, delta-5 desaturase, delta-6 desaturase, delta-15 desaturase, delta-17 desaturase, 18-22C elongase or 20-24C elongase) required for making one or more polyunsaturated fatty acid having 20 or more carbon atoms and 4 or more carbon-carbon double bonds, and the total fatty acid profile comprises 2% or more of 1 or more polyunsaturated fatty acid having 20 or more carbon atoms and 4 or more carbon-carbon double bonds; (4) an oilseed plant (V) comprising (IV); (5) seeds (VI) obtained from (V); (6) oil (VII) obtained from (V); (7) a food product or food analogue (VIII) comprising (II), (III), (VI) or (VII); (8) a beverage comprising (II) or (VII); (9) infant formula comprising (III) or (VII); (10) a nutritional supplement comprising (III) or (VII); (11) a pet food comprising (II), (III), (VI) or (VII); (12) animal feed comprising (II) or (VII); (13) a whole bean soy product comprising (II) or (VII); and (14) an aquaculture food product comprising (II), (III), (VI) or (VII). (I) is useful for producing mature seeds in which the total seed fatty acid

CC profile comprises polyunsaturated fatty acids. (IV) is useful for producing an oilseed plant having an altered fatty acid profile which involves transforming a plant with (IV), growing the transformed plant and selecting those plants, where the total fatty acid profile comprises 1% or more of one or more polyunsaturated fatty acid having 20 or more carbon atoms and 5 or more carbon-carbon double bonds. (I) enables the production of very long chain body unsaturated fatty acids. The present sequence represents a soybean annexin promoter nucleotide sequence, which is used in the exemplification of the present invention.

XX  
SQ Sequence 2012 BP; 695 A; 326 C; 213 G; 778 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2012; DB 13; Length 2012;  
Best Local Similarity 100.0%; Pred. No. 3.5e-225;  
Matches 2012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATCTTAGGCCCTTCATTATATGTTAGTGGATTGATGATTCACATGCAAGTTTATTTCAT 60  
Db ATCTTAGGCCCTTCATTATATGTTAGTGGATTGATGATTCACATGCAAGTTTATTTCAT 60  
QY 61 CCTTTTCTTTGTAATACTGACCAAGAACCAAGAAAAAAGAAAAAGGATCA 120  
Db CCTTTTCTTTGTAATACTGACCAAGAACCAAGAAAAAAGAAAAAGGATCA 120  
QY 121 TTTTGAAGGATATTTTTCGCTCTCTATTCAAATCTGTTATTTTACCAAAAAACTGTAT 180  
Db TTTTGAAGGATATTTTTCGCTCTCTATTCAAATCTGTTATTTTACCAAAAAACTGTAT 180  
QY 181 TTTTCTTACACTCTCAAGCTTTTTCGCTTCGACTCTCATGATTTTCTTCATATGCC 240  
Db TTTTCTTACACTCTCAAGCTTTTTCGCTTCGACTCTCATGATTTTCTTCATATGCC 240  
QY 241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTTGCAAAAGCTTTGTCA 300  
Db AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTTGCAAAAGCTTTGTCA 300  
QY 301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCGCTTCAAAATTTATCTGCAACCCCTAGC 360  
Db AGCTTGCAATGTACAAATTAATGTTTTTCATGCGCTTCAAAATTTATCTGCAACCCCTAGC 360  
QY 361 TATTAACTCAACATCTAAGTAAGGCTAGTGAATTTTTCGAATAGTCAAGTGCATTA 420  
Db TATTAACTCAACATCTAAGTAAGGCTAGTGAATTTTTCGAATAGTCAAGTGCATTA 420  
QY 421 ATTTCCCGTGACTATTTTGGCTTTTGACTCCAAACACTGGCCCCGTACATCCGTCCTCAT 480  
Db ATTTCCCGTGACTATTTTGGCTTTTGACTCCAAACACTGGCCCCGTACATCCGTCCTCAT 480  
QY 481 TACATGAAAAGAAATATGTTTATATTTCTTAAATTAATAATTTGTCCTTCTAAATTTTC 540  
Db TACATGAAAAGAAATATGTTTATATTTCTTAAATTAATAATTTGTCCTTCTAAATTTTC 540  
QY 541 ATATAGTTAAATATATATATATATTTTCTATCTATCTATCTATCTATCTATCTATCTAT 600  
Db ATATAGTTAAATATATATATATATTTTCTATCTATCTATCTATCTATCTATCTATCTAT 600  
QY 601 ATTTATGCATATGTAAAGTACATTTATTTTGTCTATATCTTAAATATTTTCTAAATTTAT 660  
Db ATTTATGCATATGTAAAGTACATTTATTTTGTCTATATCTTAAATATTTTCTAAATTTAT 660  
QY 661 TAAAAAAGACTGATGAAAAAATTTATCTTTTAAAGCTATATCATTTTATATATAT 720  
Db TAAAAAAGACTGATGAAAAAATTTATCTTTTAAAGCTATATCATTTTATATATAT 720  
QY 721 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 780  
Db TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 780  
QY 781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTGTGTAACA 840  
Db TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTGTGTAACA 840  
QY 841 AATCATATCATTTATGATTTGAAAGAGAGAAATTTGACAGTGAGTAATAGTGATGAGAAA 900

Db 841 AATCATATCATTTATGATTTGAAGAGAGAGAAATTCACAGTGAATTAAGTATGAGAAA 900  
Qy 901 AAATGTTGTTATTCCTTAATAAAAAACCTTAACAACATGATCTACTCTCTATTTTCATCTA 960  
Db 901 AAATGTTGTTATTCCTTAATAAAAAACCTTAACAACATGATCTACTCTCTATTTTCATCTA 960  
Qy 961 TCTCTCATTTTCATTTTCTCTTTATCTCTTTTCTTTTATTTTATCATATCATTTTCACAT 1020  
Db 961 TCTCTCATTTTCATTTTCTCTTTATCTCTTTTCTTTTATTTTATCATATCATTTTCACAT 1020  
Qy 1021 TAATTAATTTTACCTCTCTTTATTTTCTCTCTATCCCTCTCTTTATTTTCCACTCATATAT 1080  
Db 1021 TAATTAATTTTACCTCTCTTTATTTTCTCTCTATCCCTCTCTTTATTTTCCACTCATATAT 1080  
Qy 1081 ACATCCAAATTCGGGATGCTTTTATCATCTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140  
Db 1081 ACATCCAAATTCGGGATGCTTTTATCATCTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140  
Qy 1141 ATGAACTGAAAGCATTTGGCAAGTCTCTCCCTCTCTCAAGTGAATTTCCAACTCAGCAT 1200  
Db 1141 ATGAACTGAAAGCATTTGGCAAGTCTCTCCCTCTCTCAAGTGAATTTCCAACTCAGCAT 1200  
Qy 1201 TGGCATCTGATTTGATTCAGTATATCTATTGTCATGTGTAAAGTCTTTCCAACTACATAA 1260  
Db 1201 TGGCATCTGATTTGATTCAGTATATCTATTGTCATGTGTAAAGTCTTTCCAACTACATAA 1260  
Qy 1261 CTATTAATTAATCTTAATAATAAAGGATTAATATTTTCTCTCTCAATAAATTTAA 1320  
Db 1261 CTATTAATTAATCTTAATAATAAAGGATTAATATTTTCTCTCTCAATAAATTTAA 1320  
Qy 1321 ATATGTTATTTTGTGTTAGATGTATATTCGAATAAATCTAAATATATGATTAATGATTTT 1380  
Db 1321 ATATGTTATTTTGTGTTAGATGTATATTCGAATAAATCTAAATATATGATTAATGATTTT 1380  
Qy 1381 TTATATGATTAATAACATATATCAATATTAATAATGATATTTTATATATAGTGTGACA 1440  
Db 1381 TTATATGATTAATAACATATATCAATATTAATAATGATATTTTATATATAGTGTGACA 1440  
Qy 1441 CATAAATTTTAAAGGATTAATAATATGATAAATAAATTTTAAATATTTTATATTTAC 1500  
Db 1441 CATAAATTTTAAAGGATTAATAATATGATAAATAAATTTTAAATATTTTATATTTAC 1500  
Qy 1501 GAGAAAAAATATTTTAGCATAAATAAATGACAGCATATTTTCAACCTTAGTAAT 1560  
Db 1501 GAGAAAAAATATTTTAGCATAAATAAATGACAGCATATTTTCAACCTTAGTAAT 1560  
Qy 1561 TCATAAATTCCTATATGATATTTTGAATTAATAAACAAGTAATCGTTAAGGGAAGGAAATC 1620  
Db 1561 TCATAAATTCCTATATGATATTTTGAATTAATAAACAAGTAATCGTTAAGGGAAGGAAATC 1620  
Qy 1621 CTAGTCATCTCTTGCCATTTGTTTTCATGCAACAGAAAGGACGAAAAACCACTCA 1680  
Db 1621 CTAGTCATCTCTTGCCATTTGTTTTCATGCAACAGAAAGGACGAAAAACCACTCA 1680  
Qy 1681 CCATGAATCACTCTTACACCATTTTACTAGCAACAAGTCTCAACACTGAAGGCAGC 1740  
Db 1681 CCATGAATCACTCTTACACCATTTTACTAGCAACAAGTCTCAACACTGAAGGCAGC 1740  
Qy 1741 TCTCTTTCGGTTCTTTTACAACTTTCTTTGAAATAGTAGTATTTTTCATCATG 1800  
Db 1741 TCTCTTTCGGTTCTTTTACAACTTTCTTTGAAATAGTAGTATTTTTCATCATG 1800  
Qy 1801 ATTTTATTAACGTGCAAAAGATGCTTATGATAGTGCACATTTGATGACTACTA 1860  
Db 1801 ATTTTATTAACGTGCAAAAGATGCTTATGATAGTGCACATTTGATGACTACTA 1860  
Qy 1861 ATTAGAATCAATGAAAAAGCATTTGTTCTAACACGATAATCTCTGTAAGGCGTTAACTCCAAA 1920  
Db 1861 ATTAGAATCAATGAAAAAGCATTTGTTCTAACACGATAATCTCTGTAAGGCGTTAACTCCAAA 1920  
Qy 1921 GATCCAAATTTTCATATATAATTTGTGACGAAAGCAAAATGAATTCATAGCTGAGAG 1980

Db 1921 GATCCAAATTTTCATATATAATTTGTGACGAAAGCAAAATGAATTCATAGCTGAGAG 1980  
Qy 1981 AAAGGAAAGGTTAACTAAAGAACCAATACTTCA 2012  
Db 1981 AAAGGAAAGGTTAACTAAAGAACCAATACTTCA 2012  
RESULT 3  
ADRL6027  
ID ADR16027 standard; DNA; 1883 BP.  
XX  
XX ADR16027;  
XX AC  
XX XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Soybean annexin promoter truncated sequence #1.  
XX  
XX Soybean; annexin promoter; ds; dicotyledonous plant; plant.  
XX  
XX Glycine max.  
XX  
XX US2004158052-A1.  
XX  
XX 12-AUG-2004.  
XX  
XX 11-FEB-2004; 2004US-00776889.  
XX  
XX 12-FEB-2003; 2003US-0446833P.  
XX  
XX (KINN/) KINNEY A J.  
XX  
XX (LIUZ/) LIU Z.  
XX  
XX Kinney AJ, Liu Z;  
XX  
XX  
DR WPI; 2004-592806/57.  
XX  
XX New isolated nucleic acids comprising seed-specific soybean annexin and  
PT P34 soybean promoters, useful for regulating expression of at least one  
PT heterologous nucleic acid fragment in plants.  
XX  
XX Claim 2; SEQ ID NO 13; 27pp; English.  
XX  
XX The invention relates to an isolated nucleic acid fragment comprising a  
CC seed-specific soybean annexin promoter or seed-specific soybean P34  
CC promoter. The invention also relates to a recombinant expression  
CC construct comprising at least one heterologous nucleic acid fragment  
CC operably linked to any one of the isolated nucleic acid fragments above,  
CC a plant comprising the recombinant expression construct in its genome and  
CC a method of regulating expression of at least one heterologous nucleotide  
CC sequence in the plant. The plants are dicotyledonous plants, preferably  
CC soybean. The method of regulating expression of at least one heterologous  
CC nucleotide sequence in a plant comprises transforming a plant cell with  
CC the recombinant expression construct, growing fertile mature plants from  
CC a transformed plant cell and selecting plants comprising a transformed  
CC plant cell expressing the heterologous nucleotide sequence.  
CC Alternatively, plants comprising a transformed plant cell expressing the  
CC heterologous nucleotide sequence during early seed development are  
CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin  
CC and P34 promoters. The promoters are useful in regulating or promoting  
CC expression of at least one heterologous nucleic acid fragment in plants.  
CC This sequence represents a soybean annexin promoter truncated form of the  
CC invention.  
XX  
SQ Sequence 1883 BP; 645 A; 307 C; 193 G; 738 T; 0 U; 0 Other;  
Query Match 93.6%; Score 1883; DB 13; Length 1883;  
Best Local Similarity 100.0%; Pred. No. 3.2e-210;  
Matches 1883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 130 GATATTTTCGCTCCCTTAACTCAATAGTATTTTACCAAAAACTGATTTTCTCTAC 189  
Db 1 GATATTTTCGCTCCCTTAACTCAATAGTATTTTACCAAAAACTGATTTTCTCTAC 60

QY 190 ACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCTTGTATATGCCAATCACTCT 249  
Db 61 ACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCTTGTATATGCCAATCACTCT 120  
QY 250 ATTTATAAATGTCATAAGTAGTGTGAACAATTTGCAAAAGCTGTGTCATCAAAAGCTTTGCAA 309  
Db 121 ATTTATAAATGTCATAAGTAGTGTGAACAATTTGCAAAAGCTGTGTCATCAAAAGCTTTGCAA 180  
QY 310 TGTACAAATTAATGTTTTTCATGCTTTTCAAAATTTATCTGCAACCCCTAGCTATTAATCT 369  
Db 181 TGTACAAATTAATGTTTTTCATGCTTTTCAAAATTTATCTGCAACCCCTAGCTATTAATCT 240  
QY 370 AACACTTAAGTAAGCTAGTGAATTTTTCGNATAGTCATGCAAGTGCAATTAATTCGCCG 429  
Db 241 AACACTTAAGTAAGCTAGTGAATTTTTCGAATAGTCATGCAAGTGCAATTAATTCGCCG 300  
QY 430 TGACTATTTTGGCTTTTGACTCAACACTGGGCCCGTACATCGTCCCTCATTAACATGAAA 489  
Db 301 TGACTATTTTGGCTTTTGACTCAACACTGGGCCCGTACATCGTCCCTCATTAACATGAAA 360  
QY 490 AGAAATATGTTTATATTTCTTAATTAATAAATAATTTGTCCTCTCTAAATTTTTCATATAGTTA 549  
Db 361 AGAAATATGTTTATATTTCTTAATTAATAAATAATTTGTCCTCTCTAAATTTTTCATATAGTTA 420  
QY 550 ATTATATATTAATTTTCTCTATTTCTATTTAGTCTATTTTCAAAATTTATTTATGCA 609  
Db 421 ATTATATATTAATTTTCTCTATTTCTATTTAGTCTATTTTCAAAATTTATTTATGCA 480  
QY 610 TATGTAAGTAGTACATTAATTTTGTCTATATATCTTAATAATTTCTAAATTTATTAATAAAG 669  
Db 481 TATGTAAGTAGTACATTAATTTTGTCTATATATCTTAATAATTTCTAAATTTATTAATAAAG 540  
QY 670 ACTGATATGAAAAATTTATTTCTTTTAAAGCTATATCAATTTTATATATATATTTTCTTTT 729  
Db 541 ACTGATATGAAAAATTTATTTCTTTTAAAGCTATATCAATTTTATATATATATTTTCTTTT 600  
QY 730 CTTTTCTTTCAATTTCTATTTCAATTTAATAAGAAATAAATTTGTAAATTTTATTTATC 789  
Db 601 CTTTTCTTTCAATTTCTATTTCAATTTAATAAGAAATAAATTTGTAAATTTTATTTATC 660  
QY 790 AATTATATAAATAATTTTACTTTTATATGTTTTTTCACATTTTGTAAACAAATCATATC 849  
Db 661 AATTATATAAATAATTTTACTTTTATATGTTTTTTCACATTTTGTAAACAAATCATATC 720  
QY 850 ATTATGATTTGAAAGAGAGAAATTTGACAGTGAGTAATAAGTGATGAGAAAAAATGTGTT 909  
Db 721 ATTATGATTTGAAAGAGAGAAATTTGACAGTGAGTAATAAGTGATGAGAAAAAATGTGTT 780  
QY 910 ATTTTCTTAATAAATAAACCCTAAACAAACATGTATCTACTCTATTTTCATCTATCTCATTT 969  
Db 781 ATTTTCTTAATAAATAAACCCTAAACAAACATGTATCTACTCTATTTTCATCTATCTCATTT 840  
QY 970 TCATTTTCTCTTATCTCTTTCTTTTATTTTATCATATATCATTTTACATTAATTTT 1029  
Db 841 TCATTTTCTCTTATCTCTTTCTTTTATTTTATCATATATCATTTTACATTAATTTT 900  
QY 1030 TTACTCTCTTTATTTTCTCTATCCCTCTCTTTATTTCCACTCATATATACACTCCAA 1089  
Db 901 TTACTCTCTTTATTTTCTCTATCCCTCTCTTTATTTCCACTCATATATACACTCCAA 960  
QY 1090 AATTGGGGCAGCTCTTTATATCACTACTCTATCTCTCGACTAAATCATTTAAATGAAGCTG 1149  
Db 961 AATTGGGGCAGCTCTTTATATCACTACTCTATCTCTCTCCACTAAATCATTTAAATGAAGCTG 1020  
QY 1150 AAAAGCATTTGGCAAGTCTCTCCCTCTCAAGTGATTTTCCAACTCAGCATTTGGCATCTG 1209  
Db 1021 AAAAGCATTTGGCAAGTCTCTCCCTCTCAAGTGATTTTCCAACTCAGCATTTGGCATCTG 1080  
QY 1210 ATTGATTAGTATATCTATTTGATGTGTAAAGTCTTTTCCCAATATACATACTATTAAT 1269  
Db 1081 ATTGATTAGTATATCTATTTGATGTGTAAAGTCTTTTCCCAATATACATACTATTAAT 1140  
QY 1270 AATCTTAATAATAAAGGATAAATAATTTTTTTTCTTCAATAAATAAATAATATGTTAT 1329

## RESULT 4

ADRI6028  
ID ADRI6028 standard; DNA; 1719 BP.  
XX  
AC AC  
XX ADRI6028;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Soybean annexin promoter truncated sequence #2.  
XX  
KW Soybean; annexin promoter; ds; dicotyledonous plant; plant.  
XX  
OS Glycine max.  
XX  
FN US2004158052-A1.  
XX  
PD 12-AUG-2004.  
XX  
PF 11-FEB-2004; 2004US-00776889.  
XX  
PR 12-FEB-2003; 2003US-0446833P.

Db 1141 AATCTTAATAATAAAGGATAAATAATTTTTTTTCTTCAATAAATAAATAATATGTTAT 1200  
QY 1330 TTTTGTGTTAGATGTATATTTTGAATAAATCTAAATATATATGATTAATTTTATATTGA 1389  
Db 1201 TTTTGTGTTAGATGTATATTTGGAATAAATCTAAATATATGATTAATTTTATATTGA 1260  
QY 1390 TTAACAATATAATCAATTAATAATATGATATTTTTTTTATATAGTGTGACATAATTTT 1449  
Db 1261 TTAACAATATAATCAATTAATAATATGATATTTTTTTTATATAGTGTGACATAATTTT 1320  
QY 1450 ATAGGATAAATAATATGATAAATAAATTTTAAATATTTTATATTTACGAGAAAAA 1509  
Db 1321 ATAAGGATAAATAATATGATAAATAAATTTTAAATATTTTATATTTACGAGAAAAA 1380  
QY 1510 AAATATTTTACGCCATAAATAAATGACCAAGCATATTTTACAACTTAGTAAATTCATAAAT 1569  
Db 1381 AAATATTTTACGCCATAAATAAATGACCAAGCATATTTTACAACTTAGTAAATTCATAAAT 1440  
QY 1570 CCTATATGTATATTTGAAATTTAAACACAGATAATTCGTTAAGGGAAGGAATCCTACGTCAT 1629  
Db 1441 CCTATATGTATATTTGAAATTTAAACACAGATAATTCGTTAAGGGAAGGAATCCTACGTCAT 1500  
QY 1630 CTCTTGCCATTTGTTTTTTCATGCAACAGAAAGGACGAAAAACCACTCACCATGAATC 1689  
Db 1501 CTCTTGCCATTTGTTTTTTCATGCAACAGAAAGGACGAAAAACCACTCACCATGAATC 1560  
QY 1690 ACTCTTCACACCACTTTTCTAGCAAACTCTCAACAACTGAAGCCAGCTCTCTTTCC 1749  
Db 1561 ACTCTTCACACCACTTTTCTAGCAAACTCTCAACAACTGAAGCCAGCTCTCTTTCC 1620  
QY 1750 GTTCTTTTACACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTA 1809  
Db 1621 GTTCTTTTACACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTA 1680  
QY 1810 CGTGCCAAAGAGCTTTATTCATAGTAGTGACATTTGTATGTACTACTACTAATTAGACA 1869  
Db 1681 CGTGCCAAAGAGCTTTATTCATAGTAGTGACATTTGTATGTACTACTACTAATTAGACA 1740  
QY 1870 TGAATAAGCATTTGTTTAAACACGATAATCTCTGTAAGCGTTAACTCCAAAGATCCCAAT 1929  
Db 1741 TGAATAAGCATTTGTTTAAACACGATAATCTCTGTAAGCGTTAACTCCAAAGATCCCAAT 1800  
QY 1930 TCATATATAAATTTGTGACGAAAGCAAAATGAATTCATAGCTGAGAGAAAGGAAG 1989  
Db 1801 TCATATATAAATTTGTGACGAAAGCAAAATGAATTCATAGCTGAGAGAAAGGAAG 1860  
QY 1990 GTTAACTAAGAACCAATCTTCA 2012.  
Db 1861 GTTAACTAAGAACCAATCTTCA 1883

XX	(KINN//) KINNEY A J.	Qy	774	TAAATTTTATTTATCAATTTATAAAAAATATTTTACTTTATATATGTTTTTTCACATTTTGG	833
PA	(LIUZ//) LIU Z.	Db	481	TAAATTTTATTTATCAATTTATAAAAAATATTTTACTTTATATATGTTTTTTCACATTTTGG	540
XX					
PI	Kinney AJ, Liu Z;	Qy	834	TTAAACAAATCATATCATTTATGATTTGAAAGAGAGGAAATTTGACAGTGAGTAATAAGTGAT	893
XX		Db	541	TTAAACAAATCATATCATTTATGATTTGAAAGAGAGGAAATTTGACAGTGAGTAATAAGTGAT	600
DR	WPI; 2004-592806/57.				
XX					
XX	New isolated nucleic acids comprising seed-specific soybean annexin and	Qy	894	GACAAAAAATGTTTATTTTCCCTAAAAAACCCTTAAACAAACATGATCTCTACTCTCTATT	953
PT	P34 soybean promoters, useful for regulating expression of at least one	Db	601	GAGAAAAAATGTTTATTTTCCCTAAAAAACCCTTAAACAAACATGATCTCTACTCTCTATT	660
PT	heterologous nucleic acid fragment in plants.				
XX					
PS	Claim 2; SEQ ID NO 14; 27pp; English.	Qy	954	TCATCTATCTCTCAATTTTCATTTTCTCTTTATCTCTTTTATCTTTTATTTTATCATATCAT	1013
XX		Db	661	TCATCTATCTCTCAATTTTCATTTTCTCTTTATCTCTTTTATCTTTTATTTTATCATATCAT	720
CC	The invention relates to an isolated nucleic acid fragment comprising a				
CC	seed-specific soybean annexin promoter or seed-specific soybean P34	Qy	1014	TTCAATTAATTAATTTTATCTCTTTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1073
CC	promoter. The invention also relates to a recombinant expression	Db	721	TTCAATTAATTAATTTTATCTCTTTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	780
CC	construct comprising at least one heterologous nucleic acid fragment				
CC	operably linked to any one of the isolated nucleic acid fragments above,	Qy	1074	CATATATACACTCCAAATTTGGGGCATGCTTTTATCTACTCTCTCTCTCTCTCTCTCTCT	1133
CC	a plant comprising the recombinant expression construct in its genome and	Db	781	CATATATACACTCCAAATTTGGGGCATGCTTTTATCTACTCTCTCTCTCTCTCTCTCTCT	840
CC	a method of regulating expression of at least one heterologous nucleotide				
CC	sequence in the plant. The plants are dicotyledonous plants, preferably	Qy	1134	CATTTAAATGAAACTGAAAGCATTTGGCAAGTCTCTCTCCCTCTCTCAAGTGATTTCCAAC	1193
CC	soybean. The method of regulating expression of at least one heterologous	Db	841	CATTTAAATGAAACTGAAAGCATTTGGCAAGTCTCTCTCCCTCTCTCAAGTGATTTCCAAC	900
CC	nucleotide sequence in a plant comprises transforming a plant cell with				
CC	the recombinant expression construct, growing fertile mature plants from	Qy	1194	TCAGCATTTGGCATCTGATTTGATTCAGTATATCTATTTGCAATGTTGTAAGTCTTTCCACAA	1253
CC	a transformed plant cell and selecting plants comprising a transformed	Db	901	TCAGCATTTGGCATCTGATTTGATTTGATTTGCAATGTTGTAAGTCTTTTCCACAA	960
CC	plant cell expressing the heterologous nucleotide sequence.				
CC	Alternatively, plants comprising a transformed plant cell expressing the	Qy	1254	TACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTCTCTCTCATAA	1313
CC	heterologous nucleotide sequence during early seed development are	Db	961	TACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTCTCTCTCATAA	1020
CC	selected. The nucleic acid is useful as a plant promoter, e.g. annexin	Qy	1314	AAATTAATAATATGTTATTTTGTGTTAGATGTTATTTTCGAATTAATCTAAATATATATGATA	1373
CC	and P34 promoters. The promoters are useful in regulating or promoting	Db	1021	AAATTAATAATATGTTATTTTGTGTTAGATGTTATTTTCGAATTAATCTAAATATATGATA	1080
CC	expression of at least one heterologous nucleic acid fragment in plants.				
CC	This sequence represents a soybean annexin promoter truncated form of the	Qy	1374	TGATTTTATATTTGATTTAAACATATTAATCAATATTAATATATGATATTTTATATAGG	1433
CC	invention.	Db	1081	TGATTTTATATTTGATTTAAACATATTAATCAATATTAATATATGATATTTTATATAGG	1140
XX					
SQ	Sequence 1719 BP; 600 A; 273 C; 173 G; 673 T; 0 U; 0 Other;	Qy	1434	TTGTACACATAAATTTTATAAGGATAAAAAATATGATAAAATAAAATTTTAAATATTTTAA	1493
	Query Match 85.4%; Score 1719; DB 13; Length 1719;	Db	1141	TTGTACACATAAATTTTATAAGGATAAAAAATATGATAAAATAAAATTTTAAATATTTTAA	1200
	Best Local Similarity 100.0%; Pred. No. 3.4e-191;				
	Matches 1719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Qy	1494	TATTTACGAGAAAAAATAATTTTACCCATAAATTAATTAATTAATTAATTAATTAATTAAT	1553
		Db	1201	TATTTACGAGAAAAAATAATTTTACCCATAAATTAATTAATTAATTAATTAATTAATTAAT	1260
		Qy	1554	TAGTAATTCATTAATTTCCCTATATGTTATTTGAAATTAATAAAGAGAGAGAGAGAGAG	1613
		Db	1261	TAGTAATTCATTAATTTCCCTATATGTTATTTGAAATTAATAAAGAGAGAGAGAGAGAG	1320
		Qy	1614	AGGAATCTTACGTCATCTCTTGCCATTTGTTTTTTCATGCAACAGAGAGAGAGAGAGAG	1673
		Db	1321	AGGAATCTTACGTCATCTCTTGCCATTTGTTTTTTCATGCAACAGAGAGAGAGAGAG	1380
		Qy	1674	CACCTCACCATGAATCACTCTTACACACCATTTTCTAGCAACAAAGTCTCAACAACTGA	1733
		Db	1381	CACCTCACCATGAATCACTCTTACACACCATTTTCTAGCAACAAAGTCTCAACAACTGA	1440
		Qy	1734	AGCCAGCTCTCTTCCGTTTCTTTTACCAACACTTTCTTTGAAATAGTAGTAGTAGTAGTAG	1793
		Db	1441	AGCCAGCTCTCTTCCGTTTCTTTTACCAACACTTTCTTTGAAATAGTAGTAGTAGTAGTAG	1500
		Qy	1794	TACATGATTTTATTAACGTCGCAAAAGATGCTTTATTTGAAATAGTAGTAGTAGTAGTAG	1853
		Db	1501	TACATGATTTTATTAACGTCGCAAAAGATGCTTTATTTGAAATAGTAGTAGTAGTAGTAG	1560
		Qy	1854	ACTACTAAATTTAGAACATGAAAAAGCATTTGTTCTTAAACAGCATTAATCTCTGTGAAGCGGTTAA	1913

Db 1561 ACTACTAATTAGACATGAAAGCAATTTCTTAACAGATAATCCTGTGAAGCGTTAA 1620  
Qy 1914 CTCCAAAGATCCAAATTTCACTATATAAAATTTGTGACGAAAGCAAAATGAATTCACATAGCT 1973  
Db 1621 CTCCAAAGATCCAAATTTCACTATATAAAATTTGTGACGAAAGCAAAATGAATTCACATAGCT 1680  
Qy 1974 GAGAGAGAAAGGAAGGTTAACTAAGAGCAATACTTCA 2012  
Db 1681 GAGAGAGAAAGGAAGGTTAACTAAGAGCAATACTTCA 1719

RESULT 5  
ADRI16029  
ID ADR16029 standard; DNA; 1553 BP.  
XX  
AC ADR16029;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Soybean annexin promoter truncated sequence #3.  
XX  
DE Soybean; annexin promoter; ds; dicotyledonous plant; plant.  
XX  
KW Glycine max.  
XX  
OS US2004158052-A1.  
XX  
PN 12-AUG-2004.  
XX  
PD 11-FEB-2004; 2004US-00776889.  
XX  
PR 12-FEB-2003; 2003US-046833P.  
XX  
PA (KINNY/) KINNEY A J.  
XX  
PA (LIU/) LIU Z.  
XX  
PI Kinney AJ, Liu Z;  
XX  
DR WPI; 2004-592806/57.  
XX

New isolated nucleic acids comprising seed-specific soybean annexin and P34 soybean promoters, useful for regulating expression of at least one heterologous nucleic acid fragment in plants.

Claim 2; SEQ ID NO 15; 27pp; English.

The invention relates to an isolated nucleic acid fragment comprising a seed-specific soybean annexin promoter or seed-specific soybean P34 promoter. The invention also relates to a recombinant expression construct comprising at least one heterologous nucleic acid fragment operably linked to any one of the isolated nucleic acid fragments above, a plant comprising the recombinant expression construct in its genome and a method of regulating expression of at least one heterologous nucleotide sequence in the plant. The plants are dicotyledonous plants, preferably soybean. The method of regulating expression of at least one heterologous nucleotide sequence in a plant comprises transforming a plant cell with the recombinant expression construct, growing fertile mature plants from a transformed plant cell and selecting plants comprising a transformed plant cell expressing the heterologous nucleotide sequence. Alternatively, plants comprising a transformed plant cell expressing the heterologous nucleotide sequence during early seed development are selected. The nucleic acid is useful as a plant promoter, e.g. annexin and P34 promoters. The promoters are useful in regulating or promoting expression of at least one heterologous nucleic acid fragment in plants. This sequence represents a soybean annexin promoter truncated form of the invention.

Sequence 1553 BP; 552 A; 237 C; 149 G; 615 T; 0 U; 0 Other;

Query Match 77.2%; Score 1553; DB 13; Length 1553;  
Best Local Similarity 100.0%; Pred. No. 6.2e-172; Indels 0; Gaps 0;  
Matches 1553; Conservative 0; Mismatches 0;

Qy 460 CCCGTCATCCGTCCTCATTTACATGAAAGAAATATTTGTTTATATTTCTTAATTAATAA 519  
Db 1 CCCGTCATCCGTCCTCATTTACATGAAAGAAATATTTGTTTATATTTCTTAATTAATAA 60  
Qy 520 TATTGTCCTCTCAAAATTTTCATATAGTTAATTAATTAATTAATTAATTTCTTCTATTCTAT 579  
Db 61 TATTGTCCTCTCAAAATTTTCATATAGTTAATTAATTAATTAATTTCTTCTATTCTAT 120  
Qy 580 TAGTTCATTTTCAAAATTTATTTATGTCATATGTAAGTACATTTATTTTGTCTATAT 639  
Db 121 TAGTTCATTTTCAAAATTTATTTATGTCATATGTAAGTACATTTATTTTGTCTATAT 180  
Qy 640 ACTTAAATATTTCTAAATTTATTAATAAGCTGATATGAAATAATTTATTTCTTTTAAAG 699  
Db 181 ACTTAAATATTTCTAAATTTATTAATAAGCTGATATGAAATAATTTATTTCTTTTAAAG 240  
Qy 700 CTATATCATTTTATATATACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 759  
Db 241 CTATATCATTTTATATATACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 300  
Qy 760 AGAAATAAAATTTTGTAAATTTTATTTATCAAAATTTATAAAATAATTTTACTTTATATGTT 819  
Db 301 AGAAATAAAATTTTGTAAATTTTATTTATCAAAATTTATAAAATAATTTTACTTTATATGTT 360  
Qy 820 TTTTCACATTTTGTGTAAGCAAAATCATATCATTTATGATTTGAAAGAGAGGAAATTTGACAGT 879  
Db 361 TTTTCACATTTTGTGTAAGCAAAATCATATCATTTATGATTTGAAAGAGAGGAAATTTGACAGT 420  
Qy 880 GAGTAATAAGTGATGAGAAAAAATGTTTATTTCTTAAATAAAACCTTAAACAACATGT 939  
Db 421 GAGTAATAAGTGATGAGAAAAAATGTTTATTTCTTAAATAAAACCTTAAACAACATGT 480  
Qy 940 ATCTACTCTCTATTTTCATCTATCTCTCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 999  
Db 481 ATCTACTCTCTATTTTCATCTATCTCTCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 540  
Qy 1000 TTTTATCATATTCATTTTTCATTTTATTTTATTTTATTTTATTTTCTTTTCTTTTCTTTTCTTT 1059  
Db 541 TTTTATCATATTCATTTTTCATTTTATTTTATTTTATTTTATTTTCTTTTCTTTTCTTTTCTTT 600  
Qy 1060 CTCTTATTTTCCACTCATATATACACTCCAAATTTGGGGCATGCTTTTATCTACTACTCTAT 1119  
Db 601 CTCTTATTTTCCACTCATATATACACTCCAAATTTGGGGCATGCTTTTATCTACTACTCTAT 660  
Qy 1120 CTCTTCCACTAAATCATTTTAAATGAAACTGAAAGCAATTTGGCAAGTCTCTCTCCCTCCTC 1179  
Db 661 CTCTTCCACTAAATCATTTTAAATGAAACTGAAAGCAATTTGGCAAGTCTCTCTCCCTCCTC 720  
Qy 1180 AAGTGATTTCCAACTCAGCATTTGGCATCTGATTTGATTCAGTATATCTATTGCGATGTGTAA 1239  
Db 721 AAGTGATTTCCAACTCAGCATTTGGCATCTGATTTGATTCAGTATATCTATTGCGATGTGTAA 780  
Qy 1240 AAGTCTTTTCCACATACATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 1299  
Db 781 AAGTCTTTTCCACATACATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 840  
Qy 1300 TTTTCTTTTCCATATAAAATTAATTAATTTTGTGTTTGTAGATGATATTTTTCGAATAATC 1359  
Db 841 TTTTCTTTTCCATATAAAATTAATTAATTTTGTGTTTGTAGATGATATTTTTCGAATAATC 900  
Qy 1360 TAAATATATGATAATGATTTTATATGATTAATAACATATAATCAATTAATTAATTAATGATA 1419  
Db 901 TAAATATATGATAATGATTTTATATGATTAATAACATATAATCAATTAATTAATTAATGATA 960  
Qy 1420 TTTTCTTTTATAGTCTGACACATAAATTTTATAGGATTAATAATTAATTAATTAATTAATTAAT 1479  
Db 961 TTTTCTTTTATAGTCTGACACATAAATTTTATAGGATTAATAATTAATTAATTAATTAATTAAT 1020  
Qy 1480 TTTTAAATATTTTATATTTTACGAGAAAAAATAATTTTGTAGCCATAAATAATGACCAGC 1539  
Db 1021 TTTTAAATATTTTATATTTTACGAGAAAAAATAATTTTGTAGCCATAAATAATGACCAGC 1080



Qy 1540 ATATTTTACACCTTAGTAATTCATAAATCCTATATGTATATTTGAAATTAATAACAG 1599  
Db 1081 ATATTTTACACCTTAGTAATTCATAAATCCTATATGTATATTTGAAATTAATAACAG 1140  
Qy 1600 TAATCGTTAAGGGAAGGAACTCCTAGTCATCTCTTGCCATTTGTTTTCATGCAACAG 1659  
Db 1141 TAATCGTTAAGGGAAGGAACTCCTAGTCATCTCTTGCCATTTGTTTTCATGCAACAG 1200  
Qy 1660 AAGGAGCAAAACCCCTCACCATGAATCACTCTTCACACCATTTTCTAGCAACAA 1719  
Db 1201 AAGGAGCAAAACCCCTCACCATGAATCACTCTTCACACCATTTTCTAGCAACAA 1260  
Qy 1720 GTCTCAACAACTGAAGCCAGCTCTCTTCCGTTCTTTTACACACCTTTCTTTGAAATA 1779  
Db 1261 GTCTCAACAACTGAAGCCAGCTCTCTTCCGTTCTTTTACACACCTTTCTTTGAAATA 1320  
Qy 1780 GTAGTATTTTTTTTTCACATGATTTATTAAGTCGCCAAAGATGCTTATTTGAATAGAGTG 1839  
Db 1321 GTAGTATTTTTTTTTCACATGATTTATTAAGTCGCCAAAGATGCTTATTTGAATAGAGTG 1380  
Qy 1840 CACATTTGTATGTACTACTAATTAAGAACATGAAAGAGCATTTGTTCTAACACGATAATCC 1899  
Db 1381 CACATTTGTATGTACTACTAATTAAGAACATGAAAGAGCATTTGTTCTAACACGATAATCC 1440  
Qy 1900 TGTGAAGCGTTAACTCCAAAGATCCAAATTTCACTATATAAAATTTGTGACGAAAGCAAAAT 1959  
Db 1441 TGTGAAGCGTTAACTCCAAAGATCCAAATTTCACTATATAAAATTTGTGACGAAAGCAAAAT 1500  
Qy 1960 GAATTCACATAGCTGAGAGAGAAAGGAGTTAACTAAGNAGCAATACTTCA 2012  
Db 1501 GAATTCACATAGCTGAGAGAGAAAGGAGTTAACTAAGNAGCAATACTTCA 1553

## RESULT 6

ADRI6030  
ID ADRI6030 standard; DNA; 1367 BP.  
XX AC  
XX AC  
XX AC  
XX 18-NOV-2004 (first entry)  
XX Soybean annexin promoter truncated sequence #4.  
DE Soybean; annexin promoter; ds; dicotyledonous plant; plant.  
XX Glycine max.  
XX OS  
XX US2004158052-A1.  
XX 12-AUG-2004.  
XX PD  
XX 11-FEB-2004; 2004US-00776889.  
XX PF  
XX 12-FEB-2003; 2003US-0446833P.  
XX PR  
XX (KINN/) KINNEY A J.  
XX PA (LIUZ/) LIU Z.  
XX XX  
XX Kinney AJ, Liu Z;  
XX WPI; 2004-592806/57.  
XX DR  
XX New isolated nucleic acids comprising seed-specific soybean annexin and  
PT P34 soybean promoters, useful for regulating expression of at least one  
PT heterologous nucleic acid fragment in plants.  
XX Claim 2; SEQ ID NO 16; 27pp; English.  
XX PS  
XX The invention relates to an isolated nucleic acid fragment comprising a  
CC seed-specific soybean annexin promoter or seed-specific soybean P34  
CC promoter. The invention also relates to a recombinant expression  
CC construct comprising at least one heterologous nucleic acid fragment  
CC operably linked to any one of the isolated nucleic acid fragments above,

CC a plant comprising the recombinant expression construct in its genome and  
CC a method of regulating expression of at least one heterologous nucleotide  
CC sequence in the plant. The plants are dicotyledonous plants, preferably  
CC soybean. The method of regulating expression of at least one heterologous  
CC nucleotide sequence in a plant comprises transforming a plant cell with  
CC the recombinant expression construct, growing fertile mature plants from  
CC a transformed plant cell and selecting plants comprising a transformed  
CC plant cell expressing the heterologous nucleotide sequence.  
CC Alternatively, plants comprising a transformed plant cell expressing the  
CC heterologous nucleotide sequence during early seed development are  
CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin  
CC and P34 promoters. The promoters are useful in regulating or promoting  
CC expression of at least one heterologous nucleic acid fragment in plants.  
CC This sequence represents a soybean annexin promoter truncated form of the  
CC invention.

XX SQ Sequence 1367 BP; 493 A; 209 C; 137 G; 528 T; 0 U; 0 Other;

Query Match 67.9%; Score 1367; DB 13; Length 1367;  
Best Local Similarity 100.0%; Pred. No. 2.4e-150;  
Matches 1367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 646 ATATTTCTAAATTTATTAATAAAGACTGATATGAAAAATTTATCTTTTAAAGCTATAT 705  
Db 1 ATATTTCTAAATTTATTAATAAAGACTGATATGAAAAATTTATCTTTTAAAGCTATAT 60  
Qy 706 CATTTTATATATATCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 765  
Db 61 CATTTTATATATATCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 120  
Qy 766 AAATTTTGTAAATTTTATTTATCAATTTATAAAAAATTTTACTTTTATATGTTTTTCA 825  
Db 121 AAATTTTGTAAATTTTATTTATCAATTTATAAAAAATTTTACTTTTATATGTTTTTCA 180  
Qy 826 CATTTTGTAAACAAATCATATCATATGATGAAAGAGAGAAATTGACAGTAA 885  
Db 181 CATTTTGTAAACAAATCATATCATATGATGAAAGAGAGAGAAATTGACAGTAA 240  
Qy 886 TAAGTATGAGAAAAAATGTTTATTTCTTAAAAAAACCTTAAACAAACATGATCTAC 945  
Db 241 TAAGTATGAGAAAAAATGTTTATTTCTTAAAAAAACCTTAAACAAACATGATCTAC 300  
Qy 946 TCTCTATTTCACTATCTCTCAATTTCTTTTCTTTCTTTTCTTTCTTTTCTTTTCTTTTAT 1005  
Db 301 TCTCTATTTCACTATCTCTCAATTTCTTTTCTTTCTTTTCTTTCTTTTCTTTTCTTTTAT 360  
Qy 1006 CATATCATTTCAATTAATTTTACTCTCTTTTCTTTTCTTTTCTTTCTTTCTTTCTTTCTTT 1065  
Db 361 CATATCATTTCAATTAATTTTACTCTCTTTTCTTTTCTTTTCTTTCTTTCTTTCTTTCTTT 420  
Qy 1066 TTTCCACTCATATATACACTCCAAAAATGGGGCATGCTTTTACTCTACTCTACTCTCTCTC 1125  
Db 421 TTTCCACTCATATATACACTCCAAAAATGGGGCATGCTTTTACTCTACTCTACTCTCTCTC 480  
Qy 1126 CACTAAATCATTTAAATGAAACTGAAAGCATTTGGCAAGTCTCCTCCCTCTCAAGTGA 1185  
Db 481 CACTAAATCATTTAAATGAAACTGAAAGCATTTGGCAAGTCTCCTCCCTCTCAAGTGA 540  
Qy 1186 TTTCCAACTCAGCATTTGGCATCTGATTTGATTCAGTATATCTATTGCAATGTTAAAGTCT 1245  
Db 541 TTTCCAACTCAGCATTTGGCATCTGATTTGATTCAGTATATCTATTGCAATGTTAAAGTCT 600  
Qy 1246 TTCCCAATACATAAATCTTTAAATTAATCTTAAATAAAGAGTAAAAATATTTTTTTTTT 1305  
Db 601 TTCCCAATACATAAATCTTTAAATTAATCTTAAATAAAGAGTAAAAATATTTTTTTTTT 660  
Qy 1306 CTTCAATAAATTAATAATATGTTTATTTTGTGTTAGATGATATTCGAATAAATCTTAATA 1365  
Db 661 CTTCAATAAATTAATAATATGTTTATTTTGTGTTAGATGATATTCGAATAAATCTTAATA 720  
Qy 1366 TATGATAATGATTTTTTTTATTTATTTATTTAAACATATAATCAATATAAATATGATATTTT 1425  
Db 721 TATGATAATGATTTTTTTTATTTATTTATTTAAACATATAATCAATATAAATATGATATTTT 780







Db 661 TTTAGTACAACTCTCAACAACTGAAGCAGCTCTCTTCCGTTCTTTTACACA 720  
Qy 1766 CTTCTTTGAAATAGTAGTATTTTTCATCATGATTTATTAACGTCGCCAAAGATGCT 1825  
Db 721 CTTCTTTGAAATAGTAGTATTTTTCATCATGATTTATTAACGTCGCCAAAGATGCT 780  
Qy 1826 TATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAAGACATGAAGAAGCATTTGTC 1885  
Db 781 TATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAAGACATGAAGAAGCATTTGTC 840  
Qy 1886 TAACACGATAATTCCTGTGAAGCGTTAACTCAAAAGATCCAATTTCACTATATAAATGTT 1945  
Db 841 TAACACGATAATTCCTGTGAAGCGTTAACTCAAAAGATCCAATTTCACTATATAAATGTT 900  
Qy 1946 GACGAAGCAAAATGAATTCATAGCTGAGAGAGAAAAGGATTAACTAAGAAGCAA 2005  
Db 901 GACGAAGCAAAATGAATTCATAGCTGAGAGAGAAAAGGATTAACTAAGAAGCAA 960  
Qy 2006 TACTTCA 2012  
Db 961 TACTTCA 967

RESULT 9  
ID ADR16033 standard; DNA; 770 BP.  
XX AC ADR16033;  
DT 18-NOV-2004 (first entry)  
XX DE Soybean annexin promoter truncated sequence #7.  
XX KW Soybean; annexin promoter; ds; dicotyledonous plant; plant.  
XX OS Glycine max.  
XX PN US2004158052-A1.  
XX PD 12-AUG-2004.  
XX PF 11-FEB-2004; 2004US-00776889.  
XX PR 12-FEB-2003; 2003US-0446833P.  
XX PA (KINN//) KINNEY A J.  
XX PA (LIUZ//) LIU Z.  
XX PI Kinney AJ, Liu Z;  
XX WP1; 2004-592806/57.  
XX PT New isolated nucleic acids comprising seed-specific soybean annexin and  
PT P34 soybean promoters, useful for regulating expression of at least one  
PT heterologous nucleic acid fragment in plants.  
XX PS Claim 2; SEQ ID NO 19; 27pp; English.  
XX CC The invention relates to an isolated nucleic acid fragment comprising a  
CC seed-specific soybean annexin promoter or seed-specific soybean P34  
CC promoter. The invention also relates to a recombinant expression  
CC construct comprising at least one heterologous nucleic acid fragment  
CC operably linked to any one of the isolated nucleic acid fragments above,  
CC a plant comprising the recombinant expression construct in its genome and  
CC a method of regulating expression of at least one heterologous nucleotide  
CC sequence in the plant. The plants are dicotyledonous plants, preferably  
CC soybean. The method of regulating expression of at least one heterologous  
CC nucleotide sequence in a plant comprises transforming a plant cell with  
CC the recombinant expression construct, growing fertile mature plants from  
CC a transformed plant cell and selecting plants comprising a transformed  
CC plant cell expressing the heterologous nucleotide sequence.  
CC Alternatively, plants comprising a transformed plant cell expressing the

CC heterologous nucleotide sequence during early seed development are  
CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin  
CC and P34 promoters. The promoters are useful in regulating or promoting  
CC expression of at least one heterologous nucleic acid fragment in plants.  
CC This sequence represents a soybean annexin promoter truncated form of the  
CC invention.  
SQ Sequence 770 BP; 303 A; 108 C; 89 G; 270 T; 0 U; 0 Other;  
Query Match 38.3%; Score 770; DB 13; Length 770;  
Best Local Similarity 100.0%; Pred. No. 4.5e-81;  
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1243 TCCTTCCACATACATACATTAATTAATCTTAAATAAATAAAGGATAAATATTTT 1302  
Db 1 TCCTTCCACATACATACATTAATTAATCTTAAATAAATAAAGGATAAATATTTT 60  
Qy 1303 TTTCTTCATAAAATTAATAATATGTTATTTTGTGTAGATGTATTCGAATAAATCTAA 1362  
Db 61 TTTCTTCATAAAATTAATAATATGTTATTTTGTGTAGATGTATTCGAATAAATCTAA 120  
Qy 1363 ATATATGATAATGATTTTATATTTGATTAACATATATCAATATTAATATGATATTT 1422  
Db 121 ATATATGATAATGATTTTATATTTGATTAACATATATCAATATTAATATGATATTT 180  
Qy 1423 TTTTATATAGTTGTACACATAATTTTATAGGATAAATAATATGATAAATAAATTTT 1482  
Db 181 TTTTATATAGTTGTACACATAATTTTATAGGATAAATAATATGATAAATAAATTTT 240  
Qy 1483 AAATATTTTATATTTACGAGAAAAAATATTTTAGCCATAAATAATGACCAAGCATA 1542  
Db 241 AAATATTTTATATTTACGAGAAAAAATATTTTAGCCATAAATAATGACCAAGCATA 300  
Qy 1543 TTTTACAACCTTAGTAATTCATAAATTCCTATATATGATATTTGAAATTAATAAAGATAA 1602  
Db 301 TTTTACAACCTTAGTAATTCATAAATTCCTATATATGATATTTGAAATTAATAAAGATAA 360  
Qy 1603 TCGTTAAGGGAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAAACAGAAAG 1662  
Db 361 TCGTTAAGGGAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAAACAGAAAG 420  
Qy 1663 GGACGNAAAACCCCTCACCATGAATCACTCTTCACACCATTTTCTACTAGCAAAACAGTC 1722  
Db 421 GGACGNAAAACCCCTCACCATGAATCACTCTTCACACCATTTTCTACTAGCAAAACAGTC 480  
Qy 1723 TCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTACAAACACTTTCTTTGAAATAGTA 1782  
Db 481 TCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTACAAACACTTTCTTTGAAATAGTA 540  
Qy 1783 GTATTTTCTTTCATGATTTTATTAACGTCGCAAAAGATGCTTATTAAGATAGAGTGAC 1842  
Db 541 GTATTTTCTTTCATGATTTTATTAACGTCGCAAAAGATGCTTATTAAGATAGAGTGAC 600  
Qy 1843 ATTTGTAATGTACTACTAATTAAGACATGAAGAAGCATTTGTTCTAACACGATAATCCTGT 1902  
Db 601 ATTTGTAATGTACTACTAATTAAGACATGAAGAAGCATTTGTTCTAACACGATAATCCTGT 660  
Qy 1903 GAAGCGTTAACTCCAAGATCCAATTTCACTATATAAATGTGACGAAACAAATGAA 1962  
Db 661 GAAGCGTTAACTCCAAGATCCAATTTCACTATATAAATGTGACGAAACAAATGAA 720  
Qy 1963 TTCAATAGCTGAGAGAGAAAGGAAAGGTTAACTTAAGAAGCAATCTTCA 2012  
Db 721 TTCAATAGCTGAGAGAGAGAAAGGAAAGGTTAACTTAAGAAGCAATCTTCA 770  
RESULT 10  
ADRI6034  
ID ADR16034 standard; DNA; 584 BP.  
XX AC ADR16034;  
XX DT 18-NOV-2004 (first entry)



CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin  
CC and P34 promoters. The promoters are useful in regulating or promoting  
CC expression of at least one heterologous nucleic acid fragment in plants.  
CC This sequence represents a soybean annexin promoter truncated form of the  
CC invention.

XX  
SQ Sequence 425 BP; 154 A; 81 C; 64 G; 126 T; 0 U; 0 Other;

Query Match 21.1%; Score 425; DB 13; Length 425;  
Best Local Similarity 100.0%; Pred. No. 5.1e-41;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1588 ATTAACAAACAGATATCGTTAGGAAGAAATCCTACGTCATCTCTTGCCATTGTTTT 1647

Db 1 ATTAACAAACAGATATCGTTAGGAAGAAATCCTACGTCATCTCTTGCCATTGTTTT 60

QY 1648 CATGCAACAGAAAGGGACGAAACACCTCACCATGAATCACCTTCACACCACTTTT 1707

Db 61 CATGCAACAGAAAGGGACGAAACACCTCACCATGAATCACCTTCACACCACTTTT 120

QY 1708 ACTAGCAACAAAGTCTCAACAACTGAAGCCAGCTCTCTTCGGTTTCTTTTACAACT 1767

Db 121 ACTAGCAACAAAGTCTCAACAACTGAAGCCAGCTCTCTTCGGTTTCTTTTACAACT 180

QY 1768 TCTTTGAAATAGTAGTATTTTTTTTTCACATGATTAATACGTCGCAAAAGATGCTTA 1827

Db 181 TCTTTGAAATAGTAGTATTTTTTTTTCACATGATTAATACGTCGCAAAAGATGCTTA 240

QY 1828 TTGAATAGATGCACATTTGTAATCTACTTAATTAAGCAATGAAGCAATGTTCTTA 1887

Db 241 TTGAATAGATGCACATTTGTAATCTACTTAATTAAGCAATGAAGCAATGTTCTTA 300

QY 1888 ACACGATTAATCTGTGAAGGGCTTAACCTCAAAAGATCCAATTTCACTATATAAATTTGTA 1947

Db 301 ACACGATTAATCTGTGAAGGGCTTAACCTCAAAAGATCCAATTTCACTATATAAATTTGTA 360

QY 1948 CGAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGTTAACTAAGAACAAATA 2007

Db 361 CGAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGTTAACTAAGAACAAATA 420

QY 2008 CTTCA 2012

Db 421 CTTCA 425

RESULT 12

ID ADR16036

XX ADR16036 standard; DNA; 174 BP.

AC ADR16036;

XX 18-NOV-2004 (first entry)

DT Soybean annexin promoter truncated sequence #10.

DE Soybean; annexin promoter; ds; dicotyledonous plant; plant.

KW Glycine max.

OS US2004158052-A1.

XX 12-AUG-2004.

XX 11-FEB-2004; 2004US-00776889.

XX 12-FEB-2003; 2003US-0446833P.

XX (KINN/) KINNEY A J.

XX (LIUZ/) LIU Z.

XX Kinney AJ, Liu Z;

XX WPI; 2004-592806/57.

XX

PT

PT

PT

XX

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Query Match

Best Local Similarity

Matches 174;

Conservative

Mismatches

Indels

Gaps

0;

QY

Db

QY

Db

QY

New isolated nucleic acids comprising seed-specific soybean annexin and  
P34 soybean promoters, useful for regulating expression of at least one  
heterologous nucleic acid fragment in plants.

Claim 2; SEQ ID NO 22; 27pp; English.

The invention relates to an isolated nucleic acid fragment comprising a  
seed-specific soybean annexin promoter or seed-specific soybean P34  
promoter. The invention also relates to a recombinant expression  
construct comprising at least one heterologous nucleic acid fragment  
operably linked to any one of the isolated nucleic acid fragments above,  
a plant comprising the recombinant expression construct in its genome and  
a method of regulating expression of at least one heterologous nucleotide  
sequence in the plant. The plants are dicotyledonous plants, preferably  
soybean. The method of regulating expression of at least one heterologous  
nucleotide sequence in a plant comprises transforming a plant cell with  
the recombinant expression construct, growing fertile mature plants from  
a transformed plant cell and selecting plants comprising a transformed  
plant cell expressing the heterologous nucleotide sequence.

Alternatively, plants comprising a transformed plant cell expressing the  
heterologous nucleotide sequence during early seed development are  
selected. The nucleic acid is useful as a plant promoter, e.g. annexin  
and P34 promoters. The promoters are useful in regulating or promoting  
expression of at least one heterologous nucleic acid fragment in plants.  
This sequence represents a soybean annexin promoter truncated form of the  
invention.

Sequence 174 BP; 72 A; 28 C; 30 G; 44 T; 0 U; 0 Other;

Query Match 8.6%; Score 174; DB 13; Length 174;

Best Local Similarity 100.0%; Pred. No. 7.6e-12;

Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1839 GCACATTTGTAATCTACTACTAATTAGACATGAAGAAAGCAATTTCTTAACAGCAATATC 1898

Db 1 GCACATTTGTAATCTACTACTAATTAGACATGAAGAAAGCAATTTCTTAACAGCAATATC 60

QY 1899 CTGTGAAGGGCTTAACCTCAAAAGATCCAATTTCACTATATAAATTTGTGACCAAGCAAAA 1958

Db 61 CTGTGAAGGGCTTAACCTCAAAAGATCCAATTTCACTATATAAATTTGTGACCAAGCAAAA 120

QY 1959 TGAATTCACATAGCTGAGAGAGAAAGGTTAACTAAGAACCAATPACTTCA 2012

Db 121 TGAATTCACATAGCTGAGAGAGAAAGGTTAACTAAGAACCAATPACTTCA 174

RESULT 13

ID ABL34155

XX ABL34155 standard; DNA; 15548 BP.

AC ABL34155;

XX 26-MAR-2002 (first entry)

DT Human immune system associated gene SEQ ID NO: 2128.

DE Human; immune system disease; cytosine methylation; antiasthmatic;

KW antiarteriosclerotic; antianemic; cytosine methylation; antiasthmatic;

KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;

KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;

KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;

KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;

KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;

KW ds.

XX Homo sapiens.

OS WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP007537.

PF

XX 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX (EPiG-) EPIGENOMICS AG.  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2002-130909/17.  
XX Nucleic acid comprising fragment of chemically modified gene, useful for  
PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.  
XX  
XX  
XX Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German.  
XX  
XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention  
XX  
XX  
SQ Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 U; 0 Other;  
Query Match 7.8%; Score 156; DB 6; Length 15548;  
Best Local Similarity 50.4%; Pred. No. 4.5e-10;  
Matches 594; Conservative 0; Mismatches 565; Indels 19; Gaps 8;  
QY 495 TATTGTTTATATCTTAATTAATAATATGTCCTCTAAATTTTCATATAGTTAAATTAT 554  
DB 11586 TATATTATTTTATATTTATTTTATATTTTATTTTATATTTTATTTTATTTTATTTT 11645  
QY 555 TATATTACTTTTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 614  
DB 11646 ATTTTATATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 11705  
QY 615 AAGTACATATATTTTGTCTATATCTTAAATATTTCTAAATTTTAAATAAGACTGA 674  
DB 11706 TATATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 11765  
QY 675 TATGAAAAATTTATCTTTTAAAGCTATATCAATTTTATATATATCTTTTCTTTCTTTT 734  
DB 11766 TTTTATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 11825  
QY 735 CTTCATTTCTATCTAAATTAAGAATAAATTTGTAATTTTATTTTATTTATCAATTT 794  
DB 11826 TTATATTTTATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTAT--TTT 11882  
QY 795 ATAAATAATTTTACTTTATATGTTTTTTCACATTTTGTGTAACAATCATATCATTTAT 854  
DB 11883 ATTTTATATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 11942  
QY 855 GATTGAAAGAGAGAAATTCGACGTGAGTAATAGTCATGAGAAAAAATGTTATTTC 914  
DB 11943 GTTATTTTATGTTATATATATTTTATTTTATATGTTATATATATTTTATTTTATATA 12002  
QY 915 CTAAAAAACCTAAACAACATGTCATCTCTATTTTCAATCTATCTCTCAATTCATT 974  
DB 12003 TGTATATATATTTTATTTTATATTTTATTTTATATGTTATTTTATAT-GTTATTTT 12061  
QY 975 TTTCTCTTTATCTCTTTTATTTTATTTTATCATCATCTTTCATTAATTTTATTTTACT 1034  
DB 12062 TATAATGTTATTTTATATGTTATTTTATATGTTATTTTATATGTTATTTTATATGTT 12121  
QY 1035 CTCTTTATTTTCTCTCTATCCCTCTCTTATTTTCCACTCATATATACATCTCCAAAATG 1094  
DB 12122 ATTTTATATGTTATATATATGTTATTTTATATGTTATATATATATATATATATA 12181  
QY 1095 GGGCATGCTTTATCACTACTCTATCTCTCCCA---CTAAATCAATTTAAATGAACCTGAA 1151

DB 12182 TGTATATATATATGTTATATATATATATGTTATATATATATATATATATATATATAT 12241  
QY 1152 AAGCATTGGCAAGTCTCTCCCTCCCTCAAGTGAATTTCCAACTCAGCATTTGGCATCTGAT 1211  
DB 12242 ATATATATGTTATATATATATATATGTTATATATATATATGTTATATATATATATATAT 12301  
QY 1212 TGATTCAGTATATCTATTGTCATGCTGTAAAGTCTTTCCCAATACATACTATTAATTA 1271  
DB 12302 ATATGTTATATATATATATGTTATATATATATATGTTATATATATATATATATATAT 12360  
QY 1272 TCTTAATAATAATAAGAGATAAATAATTTTTTTTCTTCATAAATAATTAATAATGTT-TATT 1330  
DB 12361 TGTATATATATATATGTTATATATATATGTTATATATATATATATATATATATATAT 12420  
QY 1331 TTTTGTGTTAGATGTAATTTTCGAATAAATCTAAATATATG-----ATAATGATTTTTT 1382  
DB 12421 TTATGTTATATATGTTATATATATATATATTTTATGTTATATATTTTATATATATATAT 12480  
QY 1383 ATATTGATTAAACATATAATCAATATTAATAATATGATAATTTTTTATATAGTTGTACACA 1442  
DB 12481 GTATATATTTTATATAT-ATATATTTTATGTTATATATTTTATATATATATATATAT 12539  
QY 1443 TA-ATTTTATAGGATAAATAATATGATAAATAAATAATTTTAAATAATTTTATATTTACG 1501  
DB 12540 TATATTTTATATATATATATTTTATGTTATATATTTTATATATATATATATATATAT 12599  
QY 1502 AGAAAAAATAATTTTTCGATATAAATAATGACAGCATATTTTACAACTTAGTAATT 1561  
DB 12600 TGTATATATATATATATATATATATATATTTTCGTTATATATTTTATATATATATATAT 12659  
QY 1562 CATAAATTCCTATATGTAATTTTGAAATTTAAAAACAGATAATCGTTAAGGGAAGGAATCC 1621  
DB 12660 ATTTATTTTCGTATATATATATTTTATATATATATATATATATTTTCGTATATATAT 12719  
QY 1622 TAGCTCATCTCTGCCATTTGTTTTCATGCAAAACAGA 1659  
DB 12720 TTTATATATATATATATATATATATATATTTTCGTATATATATA 12757  
RESULT 14  
ABZ10246/C  
ID ABZ10246 standard; DNA; 8056 BP.  
XX  
XX ABZ10246;  
XX AC  
XX 16-JAN-2003 (first entry)  
XX  
XX Haematopoietic cell proliferation disorder related DNA sequence #386.  
XX Human; haematopoietic cell proliferation disorder; cytostatic;  
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;  
KW cytosine methylation state; gene; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO20027272-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 26-MAR-2002; 2002WO-EP003401.  
XX  
XX 26-MAR-2001; 2001US-0278333P.  
XX  
XX (EPiG-) EPIGENOMICS AG.  
XX  
XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;  
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;  
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;  
PI Schwobe I, Ziebarth H;  
XX  
XX WPI; 2003-018942/01.  
XX  
XX Detecting and differentiating between hematopoietic cell proliferative

PT disorders, comprises contacting a target nucleic acid with a reagent that  
PT distinguishes between methylated and non-methylated CpG dinucleotides.

XX Claim 28; SEQ ID NO 386; 117pp; English.

PS The present invention describes a method for detecting and  
XX differentiating between haematopoietic cell proliferative disorders  
CC associated with at least 1 gene and/or their regulatory regions in a  
CC subject. The method comprises contacting a target nucleic acid in a  
CC biological sample obtained from the subject with at least 1 reagent,  
CC which distinguishes between methylated and non-methylated CpG  
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118  
CC represent specifically claimed nucleotide sequences from the present  
CC invention. Oligonucleotides from the present invention can be used: for  
CC differentiating between healthy haematopoietic cells and proliferative  
CC disorder haematopoietic cells; for differentiating between acute  
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for  
CC determining the cytosine methylation state and/or single nucleotide  
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder  
CC related sequences and their complements; and as primers for the  
CC amplification of haematopoietic cell proliferation disorder related DNA  
CC sequences. The nucleotide sequences from the present invention can also  
CC be used for detecting a predisposition to, differentiation between  
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of  
CC haematopoietic cell proliferative disorders. The present method enables a  
CC highly specific classification of haematopoietic cell proliferative  
CC disorders allowing for improved and informed treatment of patients

XX SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Query Match 7.5%; Score 150.2; DB 8; Length 8056;  
Best Local Similarity 45.5%; Pred. No. 2.3e-09;  
Matches 725; Conservative 0; Mismatches 858; Indels 10; Gaps 5;

QY	16	TTATATGGTGTAGATGATTCACATGCAAGTGTATTTTCAATCCCTTTTCCTTGAA	75
DB	2543	TTATTTTAAATTAATTTATTTTATTTTATTTTTCATTAATTTTAAATTTTAA	2484
QY	76	TAACTGACCAAGAACCAAGAAAAAAGAAAGGATCAATTTTGAAGGATATT	135
DB	2483	CANAATAATAAAAAATATAATAATAAAAAAATAAATTTTAAATAATAATA	2424
QY	136	TTTCCTCGTATTCAAACTACTGTATTTTACAAAAAAAGCTGATTT-TTTCCTACACTCT	194
DB	2423	ATATATTTTTTTTAAATTTTCAAAAAAATAATAAATTTTAAATAATTTATATAAATAA	2364
QY	195	CAAGCTTTTGTGTTTCGCTCTGACTCTCATGATTTTCTTCATATGCCAATCACTCTATT	254
DB	2363	AAATTTATTTTAAAAAATAATAAATTTTAAATACAAAAATTAATAATTTATTTT	2304
QY	255	TAAATGGCATAGGTAGTGTGAACAATTTGCAAGCTTGTTCATCAAAAGCTTGCAATGTAC	314
DB	2303	AAAAATAATAAATAAAAAAATAAATTTTAAATTTTAAATTTTAAATAATAA	2244
QY	315	AAATTAATGTTTTTCATGCCCTTTCAAAATTTATCGACCCCTAGCTATTAAATCTAACAT	374
DB	2243	ATTTTATTAATAAATAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	2184
QY	375	CTAAGTAAGGCTAGTGAATTTTTCGAATAGTCAGTCAGTGAATTTTCCCGTGACT	434
DB	2183	TAAATAAATAATAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	2124
QY	435	ATTTTGGCTTTGACTCCAACTGCGCCGCTACATCCGTCCTCATTTACATCAAAAGAAA	494
DB	2123	AAAAATTAATAAATAAATTTATAAATTTATTTATTAATAATTTTAAATAATTTTAAATA	2064
QY	495	TATGTTTATATTCTTAATAAATAATTTGTCCTTTCTAAATTTTCATATATAGTTAATAT	554
DB	2063	TTTTTTAAAAAATAAAAAAATAATTTTAAATAATAATTTTAAATAATAATTTTAAATA	2007
QY	555	TATATTACTTTTCTCTATCTATTAGTTCTATTCTTCAAAATTTATTTATGATGATGT	614
DB	2006	TTATTAATAATTTTAAATAAATAAATAATAATAATAATAATAATAATAATAATAATA	1947

RESULT 15  
ABZ10246  
ID ABZ10246 standard; DNA; 8056 BP.

QY	615	AAAGTACATTATATTTTTTGGCTATATACCTTAATAATTTCTAAATTTATTAATAAGACTGA	674
DB	1946	TATTTATATATAAATTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	1887
QY	675	TATGAAAAATTTATCTCTTTTTTAAAGCTATATCATTTTATATATACCTTTTCTTTCTTTT	734
DB	1886	AAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1827
QY	735	CTTTCAATTTCTATTTCAATTTTAAATAAGAAATAAATTTTGTAAATTTTGTAAATTTT	794
DB	1826	TTTTTATTTATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA	1767
QY	795	ATAAAAAATTTTACTTTTATATGTTTTTTCACATTTTGTGTTAAACAAATCATCATAT	854
DB	1766	AAAAAATTTTATTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA	1707
QY	855	GATTGAAAGAGAGGAAATTTGACAGTGAGTAACTAGTAACTAGTAACTAGTAACTAGT	914
DB	1706	AAATTTAAATTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA	1647
QY	915	CTAAAAAAACCTAAAAACAAACATGATCTACTCTCTATTTCTCATCTCTCTCATTTCA	974
DB	1646	TTTTTAATTTTACAATTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT	1587
QY	975	TTTCTCTTTATCTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	1032
DB	1586	TAAATAATAAATTTTAAATAATAAATTTTAAATAAATAAATTTTAAATAAATAA	1527
QY	1033	CTCTCTTTATTTTCT	1092
DB	1526	ACATATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	1467
QY	1093	TGGGGCATGCGCTTTATCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1152
DB	1466	TTTATTAATAAATAAATAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT	1407
QY	1153	AGCATTTGGCAAGTCTCTCCCTCTCTCAAGTGATTTTCCAATGATTTCCAATGATTT	1212
DB	1406	AAATCAATATTTATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAA	1349
QY	1213	GATTCAGTATATCTATTTGTCATGTTGTAAGCTTTTCCCAATATACATACTATTAAT	1272
DB	1348	AAATTTATTTTATTTAAATAATTTATTAATAAATTTTATTTTATTTTATTTTATTT	1289
QY	1273	CTTAAATAAATAAAGGATAAATAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1332
DB	1288	ATTACATTTTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	1229
QY	1333	TTGTTTAGATGTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT	1392
DB	1228	TTATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA	1169
QY	1393	AACATATAATCAATTTTAAATAATGATATTTTATATATAGTTTGTAGTACATTAATTT	1452
DB	1168	AAAAATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA	1109
QY	1453	AGGATAAAAAATATGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1512
DB	1108	ACATTTTAAAAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA	1049
QY	1513	TATTTTAGCCATTAATAAATAAGCAGCATATTTT--ACAACCTTTAGTAATTTCAATA	1570
DB	1048	AAATTTCAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	989
QY	1571	CTATATCTATTTTGAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA	1603
DB	988	TAATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT	956



```

XX ABZ10246;
AC
XX
XX
XX 16-JAN-2003 (first entry)
XX
XX Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
XX Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200277272-A2.
XX
XX 03-OCT-2002.
XX
XX 26-MAR-2002; 2002WO-EP003401.
XX
XX 26-MAR-2001; 2001US-0278333P.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
XX Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu B;
XX Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
XX Schwöpe I, Ziebarth H;
XX
XX WPI; 2003-018942/01.
XX
XX Detecting and differentiating between hematopoietic cell proliferative
XX disorders, comprises contacting a target nucleic acid with a reagent that
XX distinguishes between methylated and non-methylated CpG dinucleotides.
XX
XX Claim 28; SEQ ID NO 386; 117pp; English.
XX
XX The present invention describes a method for detecting and
XX differentiating between haematopoietic cell proliferative disorders
XX associated with at least 1 gene and/or their regulatory regions in a
XX subject. The method comprises contacting a target nucleic acid in a
XX biological sample obtained from the subject with at least 1 reagent,
XX which distinguishes between methylated and non-methylated CpG
XX dinucleotides within the target nucleic acid. ABZ09861 to ABZ1118
XX invention. Oligonucleotides from the present invention can be used: for
XX differentiating between healthy haematopoietic cells and proliferative
XX disorder haematopoietic cells; for differentiating between acute
XX lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
XX determining the cytosine methylation state and/or single nucleotide
XX polymorphisms (SNPs) of haematopoietic cell proliferation disorder
XX related sequences and their complements; and as primers for the
XX amplification of haematopoietic cell proliferation disorder related DNA
XX sequences. The nucleotide sequences from the present invention can also
XX be used for detecting a predisposition to, differentiation between
XX subclasses, diagnosis, prognosis, treatment and/or monitoring of
XX haematopoietic cell proliferative disorders. The present method enables a
XX highly specific classification of haematopoietic cell proliferative
XX disorders allowing for improved and informed treatment of patients
XX
XX Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;
XX
XX Query Match 7.4%; Score 148.6; DB 8; Length 8056;
XX Best Local Similarity 47.9%; Pred. No. 3.6e-09;
XX Matches 535; Conservative 0; Mismatches 564; Indels 18; Gaps 3;
XX
XX 490 AGAATAATGTTATATCTTAATAAATAATGTCCTCTCAAAATTTTCATATAGTTA 549
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 873 ATAAAAAATTTGTAATATTTAAATTTTATATTTTAAATTTTGGTTGTTGTAATA 932
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 550 ATTATTTATTTACTTTTCTCTATTTCTATTTAGTTCTATTTTCAAAATTTATTTATGCA 609
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 933 ATTTTTTTAAAAATAAATTTTATTTTATTTTAAATGAATAAATATATATATATAAT 992
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: September 2, 2005, 17:54:23  
Job time : 787.833 secs

```

QY 610 TATGTAAGTACATTATATTTTTTGTCTATATACATAAATATTTCTAAATATTTAAAAAAG 669
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 993 TTTTTTAAAAAATAATTTATTTTAAAAAATTTTTTATTTATTTTAAATTTTGAAATTTAAAT 1052
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 670 ACTGATATGAAAAATTTATCTTTTTTAAAGCTATATCATTTTTATATATATCTTTTTCTTTT 729
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1053 ATGTAATAAATAAATAAATTTTGTGTTTAAATTTGTTTTTTTTTAAATGTTGTTATA 1112
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 730 CTTTTCTTTTCATTTTCTATTTCAATTTTAAAGAAAAATAAATTTTGTGAAAAATTTTATTTATC 789
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1113 TATTGTTTTTTTAATATTTTTTTTGATATTTTGAAAAATTTTAAAAATTAATATATTTTGAAT 1172
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 790 AATTTATAAAAAATTTTACTTTATATGTTTTTTTTCACATTTTTTGTGTAACAAATCATATC 849
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1173 ATTATTAATTTATTTTAAAAATTTTAAATATTTATTTTAAATTTTAAATATAAATA 1232
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 850 ATTATGATTGAAAGAGAGGAAATTTGACAGTGAGTAATAAGTGATGAGAAAAAATGTTT 909
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1233 ATTTTTTTGAATAAAAAAATAATGTTA- - - -TGTAATTAATTAATATAAATAATGTAAT 1288
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 910 ATTTCTTAAAAAAACCTAAACAAACATGTATCTCTCTCTATTTTCTATCTCTCTCAT 969
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1289 ATTTTTTAAAAATTAATAAATTAATAAATAATTTATTAATTTTTTAAATAAATAAT 1348
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 970 TCATTTTTCTTTATCTCTTTCTTTTATTTTTTATCATATCATTTTACATTAATTTATTT 1029
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1349 TAAATTTTAAATTTTTTATTAATAAATAATTTATTTTTTTTTTATAAATAATAATGATTTTT 1408
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1030 TTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTTCCACTCATATATACACTCCAA 1089
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1409 TTTTATTTTTTTTATTTTTTTTTTAAAAAATAAATAAATAATTTATTTTTTAAATAAAT 1468
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1090 AATTGGGGCATGCTTTATCACCTACTCTCTCTCTCCACTAAATCATTTAAATGAAACTG 1149
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1469 TATTTAAATAATTAATTTATTAATAATTAATTAATAATAAATAATAAATATGTTA 1528
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1150 AAAAGCAATGGCAAGTCTCCTCCCTCTCAAGTGATTTTCCAACTCAGCAATGGCACTCG 1209
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1529 AAAAAATTTTAAATAAATAATTTATTTTTTAAAAAATAAATAATTTTAAATTTTAAAT 1588
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1210 ATTGATTCAGTATCTATTTGCGATGTAAGAGTCTTTCCACATACATACTAATTAAT 1269
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1589 TTATATTTTATAATTTTAAATTTTAAAAAATTTGAAATAAATGAAATTTGTAATAAATA 1648
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1270 AATCTTAAATAAATAAAGGATAAATAATTTTTTTTTCTTTCATAAATAAATAATATGTTAT 1329
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1649 AATTATTTTAAATTAATATGATAAATAATTTTATTTTAAATAAATAAATAATTAAT 1708
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1330 TTT--TTGTTTAGATGATATATTGGAATAAATCTAAATATATGATAATGATTTTATATTT 1387
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1709 TATAATAATATATATTTTATAATTTTAAATTTTAAAAAATAAATAAATAAATAAATA 1768
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1388 GATTAAACATATAATCAATTAATAATATGATATTTTTTTTATATAGTTGTACACATAAT 1447
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1769 AATTAAAAAATAAATAAATTTAATTTGTTTATATATTTTAAATTTAATAAATAAATAAAT 1828
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1448 TTATAGGATAAATAAATAATGATAAATAAATAAATTTTA- - - - -AATATTTTATA 1495
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1829 TTGTTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1888
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1496 TTTTACAGAAAAAATAAATAATTTTAGCCATAAATAAATGACAGCATATTTTACACCTTA 1555
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1889 TTTTATTTAAAAAATAAATAAATTTTATAAATAAATAAATAAATAAATAAATAAATAAATA 1948
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1556 GTAAATTCATAAATTCCTATATGTTATTTTGAATAA 1592
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1949 AAAAAATATAATTTTTTTAAAAATAAATAATTAATTAATTAATA 1985
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2005, 10:34:39 ; Search time 4923.25 Seconds  
(without alignments)  
15555.832 Million cell updates/sec

Title: US-10-776-889-1  
Perfect score: 2012  
Sequence: 1 attctaggcccttgattata.....aactaagaagcaatacttca 2012

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsel:\*  
9: gb\_gsel2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	172.4	8.6	1626	7 CF238805	CF238805 AGENCOURT
c 2	165	8.2	1348	9 CG749499	CG749499 P043-4-A0
c 3	159.8	7.9	1608	9 CL118721	CL118721 ISB1-72J8
c 4	159.2	7.9	1242	9 CL068807	CL068807 CH216-115
c 5	156	7.8	1542	9 AG386981	AG386981 Mus muscu
c 6	155.8	7.7	1896	9 AG753083	AG753083 P048-1-C0
c 7	155.4	7.7	1378	9 AG350209	AG350209 Mus muscu
c 8	154.4	7.7	1210	9 CG749728	CG749728 P044-1-C0
c 9	153	7.6	1805	9 CL080711	CL080711 CH216-159
c 10	152.8	7.6	1392	9 CG757503	CG757503 P052-4-C0
c 11	152.4	7.6	1632	9 CL082569	CL082569 CH216-167
c 12	152	7.6	1594	9 CL110653	CL110653 ISB1-53P2
c 13	150.6	7.5	1594	9 CL038406	CL038406 CH216-46A
c 14	149.6	7.4	1981	9 CL082000	CL082000 CH216-165
c 15	149	7.4	1074	8 BZ696936	BZ696936 SP_Ba009
c 16	149	7.4	1391	9 CG754863	CG754863 P050-2-G0
c 17	148.8	7.4	1241	9 AG448181	AG448181 Mus muscu
c 18	148.4	7.4	1539	9 AG340947	AG340947 Mus muscu
c 19	148.4	7.4	1599	9 CL083840	CL083840 ISB1-2H14
c 20	148	7.4	2157	9 CL081966	CL081966 CH216-165
c 21	147.6	7.3	2087	9 AG333887	AG333887 Mus muscu
c 22	145.6	7.2	1745	9 AG338221	AG338221 Mus muscu
c 23	145.2	7.2	1021	9 CNS014DY	AL104032 Drosophil
c 24	145.2	7.2	1227	9 AG430010	AG430010 Mus muscu

c 25	144.8	7.2	1260	9 CL491610	CL491610 SAIL 559
c 26	144	7.2	1217	9 CL062848	CL062848 CH216-98N
c 27	144	7.2	1354	9 CG744717	CG744717 P037-2-F0
c 28	143.8	7.1	1489	9 AG350139	AG350139 Mus muscu
c 29	143.4	7.1	1277	8 CC253231	CC253231 CH261-180
c 30	143.4	7.1	1313	7 CK997149	CK997149 ip15c02_b
c 31	143.4	7.1	1454	9 CG747614	CG747614 P041-2-A0
c 32	142.6	7.1	1268	9 AG347098	AG347098 Mus muscu
c 33	142.4	7.1	1202	8 CC262481	CC262481 CH261-167
c 34	142.4	7.1	1224	9 CL077121	CL077121 CH216-143
c 35	141.6	7.0	1503	9 AG346761	AG346761 Mus muscu
c 36	140.2	7.0	1205	9 CL143963	CL143963 ISB1-1230
c 37	140.2	7.0	1928	9 CL073845	CL073845 CH216-130
c 38	140	7.0	1536	9 CL078538	CL078538 CH216-151
c 39	139	6.9	1355	9 AG346348	AG346348 Mus muscu
c 40	138.4	6.9	1493	9 CL078589	CL078589 CH216-151
c 41	137.6	6.8	1507	9 AG346189	AG346189 Mus muscu
c 42	137.2	6.8	1029	9 CNS01ZGM	AL174271 Tetraodon
c 43	137.2	6.8	1453	9 AJ591978	AL174271 Tetraodon
c 44	137	6.8	888	8 AZ549422	AZ549422 ENTDP49TF
c 45	137	6.8	1211	9 AG349657	AG349657 Mus muscu

ALIGNMENTS

RESULT 1  
CF238805/c  
LOCUS CF238805 1626 bp mRNA linear EST 05-AUG-2003  
DEFINITION AGENCOURT\_15099447 NICHD\_XGC\_Emb6 Xenopus tropicalis CDNA clone  
IMAGE:6995950 5', mRNA sequence.  
ACCESSION CF238805  
VERSION CF238805.1 GI:33442013  
KEYWORDS EST.  
SOURCE Xenopus tropicalis (western clawed frog)  
ORGANISM Xenopus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
Xenopodinae; Xenopus; Silurana.  
REFERENCE 1 (bases 1 to 1626)  
NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Office of Cancer Genomics  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: csapbs-r@mail.nih.gov  
Tissue Procurement: Robert M. Grainger  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LIAM14680 row: 9 column: 21  
High quality sequence start: 71  
High quality sequence stop: 316.  
Location/Qualifiers  
1..1626  
/organism="Xenopus tropicalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8364"  
/clone="IMAGE:6995950"  
/tissue\_type="neurala"  
/dev\_stage="embryo, stages 14-19"  
/lab\_hosts="DH10B (phage-resistant)"  
/clone\_lib="NICHD\_XGC\_Emb6"  
/note="vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb. Constructed by Invitrogen. Note: This is a Xenopus Gene Collection (XGC) library."

ORIGIN

```
Query Match      8.6%; Score 172.4; DB 7; Length 1626;
Best Local Similarity 45.8%; Pred. No. 2e-14;
Matches 526; Conservative 0; Mismatches 617; Indels 5; Gaps 3;

Qy 495 TATTGTTATATCTCTAAATTAATAATATGTCCTCTCTAAATTTTCATATAGTTAAATTAT 554
Db 1174 TATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT 1115

Qy 555 TATATTACTTTTTCTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 614
Db 1114 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT 1055

Qy 615 AAAGTACATATATATTTTGTCTATATCTTAAATATTTCTAAATTAATAAAGAGACTGA 674
Db 1054 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT 995

Qy 675 TATGAAAAATTTATCTTTTTTAAAGCTATATCAATTTATATATATATCTTTTTCTTTCTTT 734
Db 994 TATATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 935

Qy 735 CTTCCTATTTCTATCTCAATTTAATAAGAAATAAATTTTGTAATTTTTTATTTATCAATTT 794
Db 934 TTTTTTTTTTTTTTTTAAATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTT 875

Qy 795 ATAAAAATATTTACTTTATATATGTTTTTTCACATTTTTTGTAAACAAATCATATCATAT 854
Db 874 ---AATTTTTTTNNATTTATTTTTTTTTTTTTTTTTTTTTTANTNNTTTTTTTAAANAATTT 818

Qy 855 GATTGAAAGAGAGAAATTTGACGTGAGTAAATAGTGATGAGAGAAAAAATGTGTATTC 914
Db 817 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTA 758

Qy 915 CTAAAAAARACCTAAAACAACATGATCTACTCTCTATTTTCATCTATCTCTCATTTTCATT 974
Db 757 ATTNNATAAATTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTTATTTAT 698

Qy 975 TTCTCTTTATCTCTTTTATTTTTTTTATCATATCATTTTCACATTTAAATTTATTTTACT 1034
Db 697 TTTTTTTTTTTNNATTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 638

Qy 1035 CTCTTTATTTTTCTCTATCCCTCTCTATTTTCCACATATATACATCCCAAAATG 1094
Db 637 TTTTTTTTTTTTANNANATAATATATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTNNNN 578

Qy 1095 GGGCATGCCTTTATCACTACTCTATCTCCCTCACTAAATCATTTAAATGAACCTGAAAG 1154
Db 577 NAAAAAATATTTATAAATTTTTTTTTTTTTTTTTTTTTTTNNNNNNNNNNAAAAAAT 518

Qy 1155 CATTGGCAAGTCTCCTCCCTCCTCAAGTGATTTCCCAACTCAGCATTTGGCATCTGATTGA 1214
Db 517 TAAAAATTTTTTTTTTTTTTTTTTTTTTTNNNNANAAAAAANAANAANAANAANAANA 458

Qy 1215 TTCAGTATATCTATTTGATGTGTAAGAGCTCTTCCACAATACATACTATTAATTAATCT 1274
Db 457 ATTTATNTTTTTTTTTTTTTTTTAAANAANAANAANAANAANAANAANAANAANAATTTTA 398

Qy 1275 TAAATAAATAAGGATAAATATTTTTTTTTTCTTCATAAAATTTAAATATGTTATTTTTTT 1334
Db 397 TATATATCTCTATATATATATTTTTTTTTTTNTNTNTATATATANATATNTTTTTTTTT 338

Qy 1335 GTTTAGATGTATATTCGAATAAAATCTAAATATATGATAAATGATTTTTTATATGATTAAA 1394
Db 337 TTTNTNTTTTTATATATANANANATATATATATATATATATATATATATATATATATAT 278

Qy 1395 CATATAATCAATATAAATATGATATTTTTTTTATATAGGT-TGPAACATAAATTTTATAA 1453
Db 277 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 218

Qy 1454 GGATAAABAATATGATAAABAATTTT-TAATAATTTTTTATATTTTACGAGAAAAAANA 1512
Db 217 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 158
```

```
Qy 1513 TATTTTAGCATAAATAATGACCAGCATATTTTACAACCTTAGTAATTCATAAAATTCCT 1572
Db 157 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 98

Qy 1573 ATATGATATATTTGAAATTAACACAGATAATCTTTAAGGGAAGAAATCTTACCTCATCTC 1632
Db 97 ATATATATATATATATATATATANANANANNANGNGCGGNGNNCCATTNNTCTCGG 38

Qy 1633 TTGCCATT 1640
Db 37 GANNNATT 30

RESULT 2
CG749499 1348 bp DNA linear GSS 24-OCT-2003
LOCUS P043-4-A06.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
ACCESSION CG749499
VERSION CG749499.1 GI:37970425
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
          Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
          Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1348)
AUTHORS Srinivasan,J., Sinz,W., Jessee,T., Wiggers-Perebolte,L., Jansen,K.,
        Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
        pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
        Evolutionary Biology
        Max-Planck-Institute for Developmental Biology
        Spemannstr. 37-39, Tuebingen D-72076, Germany
        Tel: 00497071601371
        Fax: 00497071601498
        Email: ralf.sommer@tuebingen.mpg.de
        Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..1348
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /notes="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."

ORIGIN

Query Match      8.2%; Score 165; DB 9; Length 1348;
Best Local Similarity 48.4%; Pred. No. 2.2e-13;
Matches 549; Conservative 0; Mismatches 571; Indels 15; Gaps 3;

Qy 465 TACATCCGTCCTCATCATGAAAGAAATATTGTTTATATCTTAATAAATAATATTC 524
Db 184 TAAATATTTTTTTTTTATATAAAATATTTTATATATATATAAATTTTTTTTTTATATA 243

Qy 525 TCCTTCTAAAATTTTCATATAGTTAATATATATATATATCTATTTTCTCTATCTTATTAGTT 584
Db 244 TTTTTTATATAATTTTTTATATATATTTTTTTTTTTTTTTTTTATTTATTTATTTATTT 303

Qy 585 CTATTTTCAAATATATATTTATGCAATGTAAGTACA-----TTATATTTTTTTCGTAT 637
Db 304 TTTTTTTTATATAAATTTAAATTTTATATATTTTTTTTTTATTTTATTTTTTTTTTTTT 363

Qy 638 ATACTTAAATATTTCTTAAATTTATTAABAAAGACTGATATGAAAAATTTATCTCTTTTAA 697
Db 364 ATTTTTTTTTTTTTTAAATATATATATATTTATTTATTTTTTTTTTATTTTTTTTTTTT 423
```

RESULT 3	CU118721	1608 bp	DNA	linear	GSS 05-JAN-2004
CU118721/c	ISB1-72JB_T7.1	ISB1 Xenopus tropicalis	genomic clone	ISB1-72JB,	
LOCUS	genomic survey sequence.				
DEFINITION	CU118721				
ACCESSION	CU118721.1	GI:40612356			
VERSION	GSS.				
KEYWORDS	Xenopus tropicalis (western clawed frog)				
SOURCE	Xenopus tropicalis				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;				

REFERENCE	Xenopodinae; Xenopus; Silurana.
AUTHORS	1 (bases 1 to 1608) Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE	A physical map of the xenopus tropicalis genome
JOURNAL	Unpublished (2003)
COMMENT	Contact: Richard K Wilson Genome Sequencing Center Washington University School of Medicine Email: submissions@wuston.wustl.edu Insert Length: 75000 Std Error: 0.00 Seq primer: T7 TAATACGACTCACTATAGGG Class: BAC ends High quality sequence start: 252 High quality sequence stop: 345.
FEATURES	Location/Qualifiers
source	1..1608 /organism="Xenopus tropicalis" /mol_type="genomic DNA" /db_xref="taxon:8364" /clone="ISB1-72J8" /Clone_lib="ISB1" /notes="vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"
ORIGIN	
Query Match	7.9%; Score 159.8; DB 9; Length 1608;
Best Local Similarity	48.6%; Pred. No. 1.1e-12;
Matches	503; Conservative 0; Mismatches 522; Indels 10; Gaps 2;
Qy	480 TTACATGAAAGAAATATCTGTTATATCTTAATAATAAAATATATGTCCTCTCTAAATTTT 539
Db	1543 TTTTATTAAATATTTTTTATTTTTTAAATATATATTTTTATAAATTTATTTTATAAAATATATTT 1484
Qy	540 CATATAGTTAAATTATATATATACTTTTTTCTCTATCTATAGTCTCATTTTCCAAATAT 599
Db	1483 TTTTTTTTTTTTAAATTTTTTATATATATTTTTTATTTCTCTTTATTTTTTATTTA 1424
Qy	600 TATTATGCATATGTAAGTACATATATTTTTTGCTATATACATACTAAATATTTCTAAATTA 659
Db	1423 TTTTTTTATTTTATAAATTTATTTTTTTTATATTTTATTTTATATTTATTTTAAATTT 1364
Qy	660 TTAATAAAGACTGATATGAAAAATTTATCTTTTTTAAAGCTATATCAATTTTATATATAC 719
Db	1363 TTTTTTTATTTATTTATTTATTTTTTTTTTTTCTAAATATATATTTTTTTTATATTTAAT 1304
Qy	720 TTTTTCCTTTCTTTCTTCCATTTCTCATTTCAATTTAAATAAGAAATAAAATTTTGTAAAT 779
Db	1303 TTTTATTAAATATAAATTTTAAATTTTTTATTTATTTTTTTTTTTTATATTTTATTTATTT 1244
Qy	780 TTTATTTATCAATTTTATAAAAAATTTTACTTTATATGTTTTTTTTCACATTTTGTAAAC 839
Db	1243 TTTTATTTTATTTTATTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTTATTT 1184
Qy	840 AAATCATPATCATTTATGATTCGAAGAGAGGAAATTTGACAGTGAGTAAATGAGTATGAGAAA 899
Db	1183 TATATTTTTTTATTTATATTTTAAATATTTTTTTATTTTTTTTAAATATTTTTTATTTATATA 1124
Qy	900 AAAATGTTTATTTCTCTAAAAAAACCTAAACAATGATCTACTCTCTATTTTCATCT 959
Db	1123 ATATTTTATTTTATTTTAAATTTTTTATTTTTTTTTTTTTTTTTTTTAA-TATTTATTTTATTT 1065
Qy	960 ATCTCTCATTTTCATTTTTCTTTTATCTCTTTCTTTATTTTTTTTTTATCATATCTTACACA 1019
Db	1064 TTTTTCCTTTTTTTTTTTTTTTTATTTTAAATTTTATTTTATTTTTTTTTTTTTTTTTTTT 1005
Qy	1020 TTAATATTTTTTACTCTCTTTATTTTTTTCCTCTATGCCCTCTCTTATTTCCACTCATATA 1079
Db	1004 TTTTTTTTTTTTTTTTATATTTAAATTTTTTTATTTTTTTTTTTTTTTTATATTTATTTTAT 945
Qy	1080 TACACTCCAAAATTTGGGGCATGCGTTTATCACTACTCTATCTCTCTCCTCACTAAATCATTTA 1139
Db	944 TATATTTTTTTTAAATTTTTTTTTTATATATATTTTTTTTTTATATATATTTTTTTTTTAAATTA 885

```
Qy 1140 AATGAACTGAAAGCATGGCAAGTCTCCTCCCTCCTCAAGTATTTCCAACTCAGCA 1199
Db 884 TTTAAATTTTTTTTATTTTATTTTATATATTTTAAATTTTTTATTTTATTT 825
Qy 1200 TTGGCATCTGATTCAGTTCAGTATCTATTCAGTGTGTAAGTCTTTCCACAATACATA 1259
Db 824 TTTTTTATTTTTTTTTTTTAAATATTTTATTTTATATTTATTT-----TTA 774
Qy 1260 ACTATTAATTAATCTTAATAATAAAGGATAAAATTTTTTTTCTTCATATAAATAA 1319
Db 773 ATTTTTTTTTTTTATATATATTTTCTTATATATTTTATTTTATATATATTTTATA 714
Qy 1320 AATATGTTATTTTGTGTAGATCTATCTGAATAAATCTAAATATATGATGATGATT 1379
Db 713 TTATATATTTTATTTATTTTATTTTTTTTTTTTATTTTATTTTTTTTTTAAATTT 654
Qy 1380 TTTATATGATTAACATATAATCAATATTAATTAATGATATTTTTTTATAGTTGTAC 1439
Db 653 TTTATTTTTTTTTTTTATTAATCTTATTAATAATTTTATTTATTTCTTTTATATATTT 594
Qy 1440 ACATAATTTTATAGGATAAATAATGATAAATAAATTTAAATATTTTATATTTA 1499
Db 593 TTTTATTTTATTTTAAAAAAATAAATGAGGTTTTTTTTTTTTTTTTTTTCTCC 534
Qy 1500 CGACAAAAAAATA 1514
Db 533 CTCCAAAAAAATA 519

RESULT 4
CL068807 1242 bp DNA linear GSS 31-DEC-2003
LOCUS CH216-115B3_Sp6.1 CH216 Xenopus tropicalis genomic clone
DEFINITION CH216-115B3, Genomic survey sequence.
ACCESSION CL068807
VERSION CL068807.1 GI:40524720
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1242)
AUTHORS Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq Primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 7
High quality sequence stop: 57.

FEATURES
source
1..1242
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-115B3"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match 7.9%; Score 159.2; DB 9; Length 1242;
Best Local Similarity 49.0%; Pred. No. 1.4e-12;
```

```
Matches 519; Conservative 0; Mismatches 528; Indels 13; Gaps 3;
Qy 477 TCATTACATCAAAAGAAATATATCTTTATATTTCTTAATTTAAAAATATGTCCCTCTTAAAT 536
Db 189 TTATTTATATTAATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 248
Qy 537 TTTTCATATAGTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 589
Db 249 TTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308
Qy 590 TTCAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 649
Db 309 AGATAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 368
Qy 650 TTCCTAAATTTTAAAAAAGACATGATATGAAAAATTTATTTCTTTTAAAGCTATATCATTT 709
Db 369 AATTAATTTAAATATATATTTTATTTTATTAATAATTTAAAAATTTATTTATTTATTTAT 428
Qy 710 TTATATATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 769
Db 429 TTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 488
Qy 770 TTTGTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 829
Db 489 TATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 548
Qy 830 TTTGTAAACAATATCATATCATTTATGATTCGAAAGAGGAGGAAATTTGACAGTGAGTAATAG 889
Db 549 TATAATTTAAATATTTATTTAAATAAATTTTATTTATTTATTTATTTATTTATTTATTTAT 608
Qy 890 TGATGAGAAAAAATGTTTATTTCTTAAAAAACCCTAAACAAACATGATCTACTCTC 949
Db 609 AATTTATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 668
Qy 950 TATTTCAATCTATCTCTCATTTTCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1009
Db 669 AATTTATATCTTATATATTTTATTAATAATTTATTTATTTATTTATTTATTTATTTATTT 728
Qy 1010 TCATTTTCACATTAATTTATTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1069
Db 729 TTTTATTTTAAATTAATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 788
Qy 1070 CACTCATATATACACTCCAAAAATTTGGGGCATGCTTTTATCCTACTCTATCTCTCCTCACT 1129
Db 789 AAATAATAATATATATAATTTAAATTTTATTTATTTATTTATTTTATTTATTTATTTATTT 847
Qy 1130 AAATCATTTTAAATGAAACCTGAAAGCATTTGGCAAGTCTCTCCCTCTCTCAAGTGATTC 1189
Db 848 ATTTTAAATTTTATATTTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTAT 907
Qy 1190 CAACTCAGCATTTGGCATCTGATTTGATTCAGTATATCTATTTGCGATGTGTAAGCTCTTTCC 1249
Db 908 TTTTATTTATTTATTTTATTTATTTTAAATTTTAAATTTTAAATTTTATTTTATTTTATTT 962
Qy 1250 ACAATACATAACTATTTAAATTTTAAATAAATAAAGGATAAATAATTTTCTTTCTTTCTTC 1309
Db 963 ATATAAAATTAATAAATAATTTATTTAAATAAATAAATAATTTTATTTATTTATTTTAAAT 1022
Qy 1310 ATAAAAATTAATAATGTTATTTTGTGTAGATGATATTCGAATAAATCTAAATATATG 1369
Db 1023 TTTAAATTTTATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTTAT 1082
Qy 1370 ATAAATCATTTTATTTATTTGATTAACATATATAATCAATATTTTAAATATGATATTTTATA 1429
Db 1083 TTTAAACATTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTATTTAT 1142
Qy 1430 TAGGTGTGACACATAATTTTATTAAGGATAAATAATATGATAAATAAATAAATAAATAATTT 1489
Db 1143 TTTGTATTTATATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1202
Qy 1490 TTTATATTTTACGAGAAAAAATAATTTTATTTAGCCATAAATA 1529
Db 1203 TAATATATATTAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 1242
```

FEATURES	source	Location/Qualifiers	Score	DB	Length
Query Match			7.8%	Score 156;	DB 9;
Best Local Similarity			47.8%;	Pred. No. 3.8e-12;	
Matches	507;	Conservative	0;	Mismatches 534;	Indels 19; Gaps 2;
Qy	492	AAATATTGTTTATATCTTAAATAAAAATATGTCCTCTAAATTTTCATATAGTTAAT	551		
Db	299	AAATAATTTATTTTATATATATATATTTTAAATTTTNTTAAATTTTATTAATTTAAT	358		
Qy	552	TATTATATTACTTTTTTCTCTATTCTATAGTTCTATTTTCAAATTTATTTATGCGATA	611		
Db	359	AAATATTTTATATTTTTTTTTTTTATATATATTTTAAATAAAATTTTTTTTTTAAATATAT	418		
Qy	612	TGTAAGACATATATTTTTTGCTATATACCTTAAATATTTCTAAATTTTAAAAAAGAC	671		
Db	419	TTTATAAATTTTTTTTATTTTAAATTTTTTATAAAATATTTTTTTTTTATATTATAAATAT	478		
Qy	672	TGATATGAAAAAATTTATCTTTTTTAAAGCTATATCATTTTATATATATCTTTTTCTTTCT	731		
Db	479	AAATAATATTTTTTTATATTTTTTATATATATTAATAATTTATTTTATATTTTTTATAT	538		
Qy	732	TTTCTTTTCATTTCTATTCAAATTTTAAAGAAATAAATTTTGTAAATTTTTTATATCAAT	791		



Query Match	7.7%	Score 155.4;	DB 9;	Length 1378;
Best Local Similarity	49.6%;	Pred. No. 4.6e-12;		
Matches 544; Conservative	0;	Mismatches 532;	Indels 20;	Gaps 5;
QY	493	AATATTGTTTATATCTCTTAATTAAAAATATTGTCCCTCTCAAAATTTTCATATAGCTAATT	552	
Db	288	AGTATTATAATATATTGTTATAATATATATATATATTTTTTTTATATAGTTTTATTGAT	347	
QY	553	ATTATATTACTTTTTTCTCTATTCTATTAGTTCATTTCCAAATATATATTTATGCGATAT	612	
Db	348	TTTATTGGATTTTTTAAAGTATTATTTTAAATTAATTTGAAAATATTTTTTTTTTTTTAT	407	
QY	613	GTAAGTAGACATTA-TATTTTTGCATATATACCTAAATATTTCTAAATATTAAAAAGAC	671	
Db	408	AATATTATTTTTTTATTATTATTAATTAATAATAATTACTCCTTTTATTATTATTAAT	467	
QY	672	TGATATGAAAAATTTATCTTTTTTAAAGCTATATCATTTTATATATATACTTTTTCTCTCT	731	
Db	468	TTTTATTATTATTATTTTTTTTTTATTATTTTTTTTAAATTTTATTTTTTTTTTTTTTTT	527	
QY	732	TTTCTTTCAATTTCTATTCAAATTTAAATAAGAAAAATAAATTTTGTAATTTTTTATTATCAA	791	
Db	528	TATTATGTTTTATTATTATATTAAATTTTATATATATTATATATATTTTTTTTA	587	
QY	792	TTTATAAAAAATATTTTACTTTATAGTTTTTTTTCATTTTTTGTGTAACAATCATATCAT	851	
Db	588	TTTTTATAAATTTTTTTTATTATATATAATAAATTTTTTTTATATATATTTTTTTTTTT	647	
QY	852	TATGATTTGAAGAGAGGAAATTGACAGTGAGTAAATAGTATGAGAAAAAAATCTGTTAT	911	
Db	648	TTATTTTTTATAATTTTTTTTTTTTATTATTATTATTATTTTTTTTATTTTTTATATAT	707	
QY	912	TTCTAAAAAAAACCTAACAAACATGTATCTACTCTCTATTTTCATCTATCTCTCATTTTC	971	
Db	708	TTTTTTATATTTTATTATTATTAATTTATTTTTTTTATAAATTTTAAATTTATTATTTTT	767	
QY	972	ATTTTTCTCTTATCTCTCTTTATTTTTTTTATGATATCATTTTCACATTAATTTTTTT	1031	
Db	768	ATTATTATATTTTTTTTTTTTTTATTATTATTATTATTATTATTATTATTATTATTATTA	827	
QY	1032	ACTCTCTTTATTTTTTCTCTATCTCCCTCTCTTATTTTCCACTCATATATACACTCCA	1091	
Db	828	TATATATTTTTTATATATTTTATTATTTTTTTTTTTTT-----TATTAATTATATATTT	878	
QY	1092	TTGGGGCATGCCTTTATCACTACTCTATCTCTCCACATAATCATTTAAAAATGAAACTGAA	1151	
Db	879	TTTTATTTTTTTATTTTTTTTTTTTATATATATTTTAAATTAAGTTATTTTTTTTTTTAT	938	
QY	1152	AAGCATTTGGCAAGTCTCCTCCCCTCTCAAGTGAATTTCCAACTCAGCATTTGGCATCTGAT	1211	
Db	939	TATTATTATTATTTTGTTTTTTATTATATTTTTTATTATTTTTTTTATTATTATTATTTT	998	
QY	1212	TGATTCAGTATATCTATTGTCATGTGTAAAGTCTTTCCCAAAATACATAACTATTAAATAA	1271	
Db	999	TTATTTTTTTATATATTTTAAATATATTTTAAATTTTAAATTTTTTTTATATTTTTTTT	1058	
QY	1272	TCTTAAATAAATAAGGATAAAATATTTTTTTTTTCTTCATAAAAATAAAAATATGTTATTTT	1331	
Db	1059	TTTTTAAATATAT-----TTAAATTTTATTATATGATTTTTTATTATATATATATTA	1113	
QY	1332	TTTGTTTAGATGATATTCGAAATAAAATCTAAAAATATATGATAATGATTTTTTTATATGAT	1391	
Db	1114	TTTTTTTTTTATTTTTTATATTTTTTTATATTTTTTTTATATATTAATTAATTTTTTAT-	1172	

```
QY 675 TATGAAAAATTTATTCCTTTTAAAGCTATATCATTTTATATATACCTTTTCTTTCTTTT 734
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 387 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 446

QY 735 CTTTCATTTCTATCAATTTAATAAGAAATAAATTTTGTAAGATTTTATTTATCAATTT 794
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 447 NNITTTTNTTNTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 506

QY 795 ATAAAAATATTTACTTTATATGTTTTTTCACATTTTGTGTAACAAATCATATCATAT 854
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 507 NTNTNTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 566

QY 855 GATTGAAAGAGAGAAATTCACAGTGAGTAATAAGTCATGAGAAAAAAATGTGTTATTC 914
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 567 ABAATAANTTTAATTTATTTTATTTTAAATAATTTAATNTTAAATAATATATAA 626

QY 915 CTAAAAAAACCTTAAACAACATGATCTACTCTCTATTTCACTATCTCTCAATTCATT 974
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 627 ATTTTNAAAAAATAAATAATATTTATTTNTATTTTATAATTAATAAATAAATAAATTA 686

QY 975 TTCTCTTTATCTCTTTCTTTTATTTTATCATATCATTTTCAATTAATTTTACT 1034
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 687 TAAATAATTTTAAATTTTAAATTTTNTATTTATTAATAATTTTAAATTTTATAA 746

QY 1035 CTCTTTATTTTCTCTATCCCTCTCTCTTTTCCACTCATATATACACTCCAAAATG 1094
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 747 TATTTAATTTATTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 806

QY 1095 GGGCATGCCCTTATCACTACTCTATCTCTCCACT-----AAATCAATTTAAATGAA 1145
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 807 TTATATTTTNTAATTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 866

QY 1146 ACTGAAAGCATGGCAAGCTCCTCCCTCCCTCAAGTGATTTCCCACTCAGCATGGCA 1205
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 867 TTTAATAATAATTTTAAATTTATTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 926

QY 1206 TCTGATTTGATTCAGTATATCTATTCATGTGTAA---AAGCTTTCCACATACATAACT 1262
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 927 ATAATTTAABAAAATTTATATTTTATTTTAAATTTTATTTTATTTTATTTTATTTT 986

QY 1263 ATTAATTAATCTTAAATAAAGGATAAAATATTTTTTTTTTCTTCCATAAAATTAAT 1322
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 987 TTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 1046

QY 1323 ATGTTATTTTGTGTTAGATATTCGATAAATCTAATAATATATATATGATATGATTTTT 1382
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1047 TTTTAAAAAAATTTTAAATATAAATTTTATTTTAAATTTAAATTAATTAATTAANTT 1106

QY 1383 ATATTGATTAACATATAATCAATTAATAATATGATATTTTTTATATAGTTGTACACA 1442
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1107 ATTTAATATATTTAATAAATAATTTTAAATAATTTATTTATTTATATATTTTATATT 1166

QY 1443 TAATTTTATAGGATAAAAAATATGATAAAAAATTAATTTAAAT 1486
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1167 TATTTATATATTTTTTTTATTTTATTTTATATTAATTTATTTATTTATTTATTTATTT 1210
```

```
RESULT 9
CL080711 1805 bp DNA linear GSS 31-DEC-2003
LOCUS CH216-15918.RM4.1 CH216 Xenopus tropicalis genomic clone
DEFINITION CH216-15918, genomic survey sequence.
ACCESSION CL080711
VERSION CL080711.1 GI:40536624
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1805)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
```

TITLE Mardis,E. and Wilson,R.  
JOURNAL A physical map of the xenopus tropicalis genome  
COMMENT Unpublished (2003)  
Contact: Richard K Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu  
Insert Length: 175000 Std Error: 0.00  
Seq primer: RM4 ctcaaggcgcatcggtcgagc  
Class: BAC ends  
High quality sequence start: 690  
High quality sequence stop: 766.

## FEATURES

source

1..1805  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/strain="Nigerian frog"  
/db\_xref="taxon:8364"  
/clone="CH216-15918"  
/sex="male"  
/cell\_line="Stock 248 F7A2, inbred N7"  
/clone\_lib="CH216"  
/notes="Vector: pTABAC2.1; CHORI-216 Xenopus tropicalis  
BAC library"

## ORIGIN

Query Match 7.6%; Score 153; DB 9; Length 1805;  
Best Local Similarity 46.9%; Pred. No. 9.5e-12;  
Matches 623; Conservative 0; Mismatches 687; Indels 19; Gaps 4;  
QY 495 TATTGTTTATATCTTTAAATAAATAATTTGTCCTTTCTAAATTTTCATATAGTTAATTAT 554  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 179 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 238  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 555 TATATTTACTTTTCTCTATCTCTATTTAGTCTTATTTTCAAAATTTATTTATCATATGT 614  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 239 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 298  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 615 AAAGTACATTTATTTTCTGTATATATCTTAAATATTTCTAAATTTATTAATAAAAGACTGA 674  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 299 TTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 358  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 675 TATGAAAAATTTATCTTTTAAAGCTATATCAATTTTATATATATATATATCTTTTCTTTT 734  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 359 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 418  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 735 CTTTCAATTTCTATTTCAATTTAATAAGAAATAAATTTTGTAAATTTTATTTATCAATTT 794  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 419 TTATCAATTTTATATAATATTTATTTATGATAAACTTAAATTTATTTATTTATTTATTT 478  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 795 ATAAAAATATTTTATCTTTTATATGTTTCTTTTTCACATTTTGTAAACAAATCATATCAT 854  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 479 ATTATTTGATTTTCTTTTATTAATAATTTATTTATTTTCTNTATGTTTATATATTTT 538  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 855 GATTGAAAGAGAGAAATTCACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTC 914  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 539 ATATAATAATTTTATTTTATTTTAAATATCAATATACATAAATTTATACATATTTTACAT 598  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 915 CTAAAAAAACCTTAAACAACATGATCTCTCTATTTTCATCTCTCTCATTTTCATT 974  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 599 TATATTTATTTATTTTCATCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 658  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 975 TTTCTCTTTATCTCTTTCTTTTATTTTCTTTTATCATATCATTTTCAATTAATTTTACT 1034  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 659 TTTATTTT---TTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 715  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 1035 CTCTTTATTTTCTCTCTATCCCTCTCTTATTTTCCACTCATATATACACTCCAAAATG 1094  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 716 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 775  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 1095 GGGCATGCCCTTATCACTACTCTCTCTCCCACTAAATTAATTAAGAACTGAAAAAG 1154  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 776 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 835  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||







KEYWORDS GSS.  
SOURCE Xenopus tropicalis (western clawed frog)  
ORGANISM Xenopus tropicalis  
Eukaryota; Bacteria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
Xenopodinae; Xenopus; Silurana.  
REFERENCE 1 (bases 1 to 1594)  
AUTHORS Krematzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,  
Mardis, E. and Wilson, R.  
TITLE A physical map of the xenopus tropicalis genome  
JOURNAL Unpublished (2003)  
COMMENT Contact: Richard K Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@wustl.edu  
Insert Length: 75000 Std. Error: 0.00  
Seq primer: Sp6 ATTAGTGACACTATAG  
Class: BAC ends  
High quality sequence start: 390  
High quality sequence stop: 470.  
Location/Qualifiers  
source 1..1594  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:8364"  
/clone="ISB1-53P23"  
/clone\_lib="ISB1"  
/note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC  
Library Segment 1"

ORIGIN

Query Match 7.6%; Score 152; DB 9; Length 1594;  
Best Local Similarity 47.7%; Pred. No. 1.3e-11;  
Matches 479; Conservative 0; Mismatches 520; Indels 5; Gaps 1;

Qy 495 TATGTTTATATCTCTTAATTAAGAAATATGTCCTCTCAAAATTTTCATATAGTTAATAT 554  
Db 313 TTTTATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 372  
Qy 555 TATATTACTTTTCTCTATTCTATTAGTCTATTTTCAAATTTATTTATGCTATGTT 614  
Db 373 TTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 432  
Qy 615 AAAGTACATTTATTTTTCGTATATCTAAATATTTCTAAATTTTAAATAAGACTGA 674  
Db 433 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 492  
Qy 675 TATGAATAATTTATCTTTTAAAGCTATATCATTTTATATATATCTTTTCTTTTCTTTT 734  
Db 493 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 552  
Qy 735 CTTCATTTTCTATTTCAATTTAATAAGAAATAAAATTTTGTAATAATTTTATTTATCAATTT 794  
Db 553 TTTTATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 612  
Qy 795 ATAAAAATTTTACTTTTATATGTTTTTTCACATTTTGTGTAACAAATCATATCATAT 854  
Db 613 TTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 672  
Qy 855 GATTGAAGAGAGAAATTCGACGTGAGTAAATAGTGATGAGAAAAAATGTTGTTATTC 914  
Db 673 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 727  
Qy 915 CTAAAAAACCCTAAACAAACATGTAATCTCTCTATTTCACTATCTCTCTCATTTTCAT 974  
Db 728 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 787  
Qy 975 TTTCTCTTTATCTCTTTCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTACT 1034  
Db 788 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 847  
Qy 1035 CTCTTATTTTCTCTCTATACCTCTCTTATTTTCCACTCATATATACATCCAAAAATG 1094

Best Local Similarity 47.9%; Pred. No. 2.le-11; Matches 481; Conservative 0; Mismatches 518; Indels 5; Gaps 2;		
Qy	495 TATTGTTTATATCTTAATAAATAATATGTCCTTCTAAATTTTCATATAGTTAATTAT	554
Db	1558 TTTTITTTTTTTTATATATTTTAAATATTTTATTTATTTTNTNTNTTTTTTTTTTTT	1499
Qy	555 TATATTACTTTTTCTCTATTCTATTAGTTCATTTTCAAATATATTATTTATGCAATGT	614
Db	1498 TTTTITTATNTAT	1439
Qy	615 AAGTACATATATTTTGGTATATACCTTAAATATTTCTAAATATTTAAAAAAGACTGA	674
Db	1438 TTTTATTTATTTTTTTT---TATTTTTTTTATTTATTTTTTTTTTTTATATATTTNTT	1382
Qy	675 TATGAATAATTTATCTTTTTTAAGCTATATCATTTTATATATACTTTTTCTTTCTTTT	734
Db	1381 TTATTTATTTTTTTTTTTTTTTTATTTTATTTTTTTTTTTTTTTTTTTTTTATNTTTT	1322
Qy	735 CTITTCATTTTCTATTCAATTTTAAGAAATAAATTTTGTAAATTTTTTATTTATCAATTT	794
Db	1321 TATTTTTTTTAAATT	1262
Qy	795 ATAAAAATATTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATAT	854
Db	1261 TTTTITTTATTTAT	1202
Qy	855 GATTGAAGAGAGAAATTCACAGTGAGTAATAAGTCATGAGAAATAAATGCTGTTATTC	914
Db	1201 TATTTTATT	1142
Qy	915 CTAAAAAAACCTAAACAAACATGATCTACTCTCTATTTTCATCTATCTCTCATTTTCAT	974
Db	1141 TTTTATATTTTTTTTATTTTATTTTTTTTATTTTTTTTATTTTNTTTTTTTTTTTTTT	1082
Qy	975 TTTCTCTTTATCTCTTTTATTTTTTTTATCATATCATTTTACATTTATTTTACT	1034
Db	1081 TTTTITTTTTTATTTTTTATT	1022
Qy	1035 CTCTTTATTTTTCTCTCTATCCCTCTCTATTTTCCACTCATATATACATCCCAAAATTC	1094
Db	1021 TTTTITTAT	962
Qy	1095 GGGCATGCCCTTATCACTACTCTATCTCTCCCACTAAATCATTTTAAATGAACCTGAAAG	1154
Db	961 TTTTATTATTTAA-	903
Qy	1155 CATGGCAAGTCTCTCCCTCCCTCAAGTCGATTTCCAACTCAGCATTTGGCATCTGATGA	1214
Db	902 -TTTAT	844
Qy	1215 TTCAGTATATCTATTGCATGTAAGTAAGTCCTTTCCACAATACATACTATTAATTAATCT	1274
Db	843 TTTTITTAT	784
Qy	1275 TAAATAATAAGGATAAAAATATTTTTTTTTTCTTCATAAAATTAATAATGTTATTTTTT	1334
Db	783 TTTTATTTTATTAT	724
Qy	1335 GTTTAGATGATATTCGAATAAATCTAAATATATGATTAATGATTTTATATGATTAATA	1394
Db	723 TTTTITTTTTTATTTATTTA	664
Qy	1395 CATATAATCAATATTAATAATGATTTTTTTTTTTATATAGTTTGTACACATAATTTATAAG	1454
Db	663 TTTTITT	604
Qy	1455 GATAAAAAATATGATAAAAAATAAATTTTAAATATTTTTTATATTT 1498	
Db	603 TTTTATTTTATTTNATATTT	560

FEATURES

source

1. -1981

/organism="Xenopus tropicalis"

/mol\_type="genomic DNA"

/strain="Nigerian frog"

/db\_xref="taxon:8364"

/clone="CH216-165P18"

/sex="male"

/cell\_line="Stock 248 F7A2, inbred N7"

/clone\_lib="CH216"

/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

ORIGIN

Query Match 7.4%; Score 149.6; DB 9; Length 1981;

Best Local Similarity 48.0%; Pred. No. 2.8e-11;

Matches 527; Conservative 0; Mismatches 559; Indels 12; Gaps 3;

Qy 495 TATTGTTTATATCTTAATAAATAATATGTCCTTCTAAATTTTCATATAGTTAATTAT 554

Db 763 TTTTATTTTTTTTTTAAATTTTATTTTATTTTATGTTTATTTGTTTTTATTTTATTTT 822

Qy 555 TATATTACTTTTTCTCTATTTCTATTAGTTCATTTTCAAATTTATTTATTCATCATGT 614

Db 823 TTTTITTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTATTTTATTTTATTTGT 882

Qy 615 AAGTACATATATTTTTCATATATCTTAATAATTTCTAAATATTTAAAAAAGACTGA 674

Db 883 TATTTATTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 942

Qy 675 TATGAATAATTTATTTCTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTCTTTT 734

Db 943 TTTTITGTTATTTTTTTTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 1002

Qy 735 CTTTCAATTTTCTATTCAAATTTAATAAGAAATAAATTTTGTAAATTTTATTTATTCATTT 794

Db 1003 ATTGTATTTTTTTTTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1062

Qy 795 ATAAAAATTTTATCTTTATATGTTTTTTCACATTTTTTGTAAACAAATCATATCATAT 854

Db 1063 TTATTTATTTTTTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1122

Qy 855 GATTGAAGAGAGAAATTCACAGTGAGTAATAAGTCATGAGAAAAAATGCTGTTATTC 914

Db 1123 TTTTITATATTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1182

Qy 915 CTAAAAAAACCTAAACAAACATGATCTACTCTTATTTTCATCTATCTCTCATTTCAATT 974





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 13:28:41 ; Search time 247.78 Seconds  
(without alignments)  
13286.738 Million cell updates/sec

Title: US-10-776-889-1  
Perfect score: 2012  
Sequence: 1 atcttaggccttgattata.....aactaagaagcaactactca 2012

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	155.6	7.7	18773	4	US-09-949-016-14164 Sequence 14164, A
C 2	135.4	6.7	187169	4	US-09-949-016-12776 Sequence 12776, A
C 3	135.4	6.7	191569	4	US-09-949-016-15940 Sequence 15940, A
C 4	134	6.7	18773	4	US-09-949-016-14164 Sequence 14164, A
C 5	122.4	6.1	612	4	US-09-902-540-1357 Sequence 1357, Ap
C 6	122.4	6.1	95255	4	US-09-949-016-17067 Sequence 17067, A
C 7	114.2	5.7	119153	4	US-09-949-016-12378 Sequence 12378, A
C 8	110	5.5	29717	4	US-09-949-016-16284 Sequence 16284, A
C 9	110	5.5	60376	4	US-09-949-016-12423 Sequence 12423, A
C 10	109.4	5.4	30820	4	US-09-949-016-17145 Sequence 17145, A
C 11	105.8	5.3	147382	4	US-09-949-016-14624 Sequence 14624, A
C 12	105.6	5.2	30820	4	US-09-949-016-17145 Sequence 17145, A
C 13	105.4	5.2	134987	4	US-09-949-016-13348 Sequence 13348, A
C 14	105.4	5.2	134987	4	US-09-949-016-15349 Sequence 15349, A
C 15	105.4	5.2	134987	4	US-09-949-016-15350 Sequence 15350, A
C 16	105.4	5.2	134987	4	US-09-949-016-15507 Sequence 15507, A
C 17	105.4	5.2	134987	4	US-09-949-016-15508 Sequence 15508, A
C 18	105.4	5.2	134987	4	US-09-949-016-15509 Sequence 15509, A
C 19	104.8	5.2	55886	4	US-09-949-016-15129 Sequence 15129, A
C 20	104	5.2	1039	4	US-09-902-540-1280 Sequence 1280, Ap
C 21	103.6	5.1	95255	4	US-09-949-016-17067 Sequence 17067, A
C 22	103	5.1	119153	4	US-09-949-016-12378 Sequence 12378, A
C 23	102.6	5.1	298336	4	US-09-949-016-16300 Sequence 16300, A
C 24	102.4	5.1	19124	2	US-08-487-8268-13 Sequence 13, Appl
C 25	102.2	5.1	47781	4	US-09-949-016-16492 Sequence 16492, A
C 26	102.2	5.1	47781	4	US-09-949-016-16493 Sequence 16493, A
C 27	102.2	5.1	47781	4	US-09-949-016-16494 Sequence 16494, A

C 28	102	5.1	640681	4	US-09-730-988-1 Sequence 1, Appli
C 29	101.8	5.1	19124	2	US-08-487-8268-13 Sequence 13, Appli
C 30	101.4	5.0	47781	4	US-09-949-016-16492 Sequence 16492, A
C 31	101.4	5.0	47781	4	US-09-949-016-16493 Sequence 16493, A
C 32	101.4	5.0	47781	4	US-09-949-016-16494 Sequence 16494, A
C 33	101	5.0	12313	4	US-09-949-016-13248 Sequence 13248, A
C 34	100.6	5.0	114139	4	US-09-949-016-16536 Sequence 16536, A
C 35	99.6	5.0	25067	4	US-09-949-016-11794 Sequence 11794, A
C 36	99.6	5.0	25441	4	US-09-949-016-14232 Sequence 14232, A
C 37	97	4.8	59519	4	US-09-949-016-13504 Sequence 13504, A
C 38	96.8	4.8	19438	4	US-09-949-016-12699 Sequence 12699, A
C 39	96.4	4.8	205044	4	US-09-949-016-15851 Sequence 15851, A
C 40	96.4	4.8	205044	4	US-09-949-016-15852 Sequence 15852, A
C 41	96.4	4.8	205044	4	US-09-949-016-15853 Sequence 15853, A
C 42	96.4	4.8	223471	4	US-09-949-016-12387 Sequence 12387, A
C 43	96.4	4.8	223471	4	US-09-949-016-12724 Sequence 12724, A
C 44	96.4	4.8	223471	4	US-09-949-016-12725 Sequence 12725, A
C 45	96.2	4.8	19438	4	US-09-949-016-12699 Sequence 12699, A

ALIGNMENTS

RESULT 1  
US-09-949-016-14164/c  
; Sequence 14164, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949.016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14164  
; LENGTH: 18773  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14164

Query Match	7.7%	Score 155.6; DB 4; Length 18773;
Best Local Similarity	49.5%	Pred. No. 5.9e-18;
Matches	547; Conservative	0; Mismatches 544; Indels 15; Gaps 5;
Qy	494	ATATTGTTTATCTTAAATTAATAATTTGTCCTCTCTAAATTTTCATATAGTTAAATTA 553
Db	18193	ATAATAATTAATGATTTTATATAATAATTAATAATAATAATAATAATAATAATAATA 18134
Qy	554	TTATATTACTTTTCTCTCTATTCTATTAGTTCTCTATTTTCAAATTAATTTATGTCATG 613
Db	18133	TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 18074
Qy	614	TAAAGTCATTAATTTTGGCTATATACCTTAATAATTTCTAAATTTTAAAAAAGACTG 673
Db	18073	TTGTATATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 18014
Qy	674	ATATGAAAAATTTATCTTTTAAAGC-TATATCATTTTATATATACCTTTCTTTCTTCT 732
Db	18013	TTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17954
Qy	733	TTCTTTTCATTTTCTCAATTTAATAAGAAATAAAATTTTCTAAATTTTATTTATCAAT 792
Db	17953	TCATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17894
Qy	793	TTATAAAAAATTTTACTTTTATATGTTTTTTCACATTTTGTGTTAAACAATCATATCA 852









```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17067
; LENGTH: 95255
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(95255)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17067

Query Match      6.1%; Score 122.4; DB 4; Length 95255;
Best Local Similarity 46.1%; Pred. No. 3.6e-12;
Matches 520; Conservative 0; Mismatches 601; Indels 7; Gaps 3;

QY 488 AAAGAAATATTGTTTATATCTTAATTAATAAATATGTCCTCTCAAAATTTTCATATAGT 547
Db 43253 ATACAATTCGTATATATTTTGTATATAAATACAAATTCGTATATATTTTGTATATAAT 43194

QY 548 TAATTAATTAATTAATTTTCTCTTAATCTATTAGTCTATTTTCAAAATTAATTAATG 607
Db 43193 ATACAATTCGTATTA-TATTTGTATATAATATATAAATTCGTATATATTTTGTATACAATA 43135

QY 608 CATATGTAAGTACATATATATTTTGTCTATATCTTAAATATTTCTAAATTTATAAAAAA 667
Db 43134 TATAATTCGTATATATATTTTGTATATAAATATATAATTTGTTATATATTTTGTATATAAT 43075

QY 668 AGACTGATATGAAAAATTAATCTTTTAAAGCTATATCAATTTTATATATATCTTTTCTT 727
Db 43074 ATATAATTTACATATATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43015

QY 728 TTCTTTTCTTCATTTCTCTCAATTTAAATGAAGAAATTAATTTTGTAAATTTTATTTA 787
Db 43014 ATTAATATATAATATAATTAATATAATAAATATATATTAATTAATTAATTAATTAATAA 42955

QY 788 TCAATTTTAAAAATTAATTTTCTTATATGTTTTTTCACATTTTGTGTTAAACAATCATATA 847
Db 42954 TTCTATATATATTTTGTATATAATAATAAATCATATATTAATTAATTAATTTCTATATATG 42895

QY 848 TCATTATGATTTGAAGAGAGGAAATTTGACAGTGAGTAATAGTGATGAGAAAAAATGTG 907
Db 42894 TTATATTTTGTATATAATATAATCATATATAATATAATTAATTTCTATATATATTTTGTATA 42835

QY 908 T-TATTTCTTAAAAAAACCCTAAACAAACATGATCTACTCTCTATTTTCATCTATCTCTC 966
Db 42834 TATAATATAATAATATAATATATATAATAATCTATATATCTTTGTATATAATAATAATTC 42775

QY 967 ATTTCATTTTCTCTTAATCTTTCTTTATTTTATATATATATATATATATATATATTA 1026
Db 42774 TGTATATTTTGTATATAATAATACAAATTTCTGTATATATTAATTTTGTATATAATAACAA 42715

QY 1027 TTTTTCATCTTTTATTTTCTCTATCCCTCTCTTATTTTCCACTCATATACACCTC 1086
Db 42714 TTCGTATATTTTGTATATAATAATACAAATTCATGTTATATTTTGTATATACCATGTT 42655

QY 1087 CAAATTCGGGATGCCCTTTATCTACTCTATCTCTCCCTCACTAAATCAATTTAAATGAAA 1146
Db 42654 ATATTTTGTATATAATAATAATTTGTATATTTTGTATATAATAATAATAATTTCTGTATAT 42595

QY 1147 CTGAAAAGCATGGCAAGTCTCTCCCTCCCTCAAGTGAATTTCCAACTCAGCATTTGGCAT 1206
Db 42594 TATATTTTGTATATAATAATAATTTCTGTATATTTATATTTTGTATATAATAATAATTT 42535

QY 1207 CTGATTTGATTTCA----GTATATCTATTGGCATGTGTAAAGTCTTTCCACAATACATAAC 1261
Db 42534 CTGTTATATATATTTTGTATATAAATTTCTGTTATATATATATTTTGTATATAATAATTAAT 42475

QY 1262 TATTAATTAATCTTAAATAAATAAAGGATAAATAATTTTTTTTCTTCAAAAAATTTAAA 1321
```

---

```
Db 42474 TATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 42415
QY 1322 TATGTTATTTTGTGTTAGATGTAATTCGAATAAAATCGAATAATATATGATGAATGATTTT 1381
Db 42414 TAAATATATAAATAATATATATATATATATATATATATATATATATATATATATATATAC 42355
QY 1382 TATATTCGATTAAACATATATAATCAATATTAATATATATATATATATATATATATATATAT 1441
Db 42354 GATATATATATATATTAATATATATATATATATATATATATATATATATATATATATAT 42295
QY 1442 ATTAATTTTATAAGCATAAATAATATGATAAATAAATAATTTTAAATATTTTATATATTTACG 1501
Db 42294 ATATATTAATAAATAATATATATATATATATATATATATATATATATATATATATATATAT 42235
QY 1502 AGAAAAAATAATTTTAGCCATAAATAAATGACCAGCATATTTTACAACTTAGTAAT 1561
Db 42234 ATTATATAAATAATATATATATATATATATATATATATATATATATATATATATATATAT 42175
QY 1562 CATAAATTCCTATATATGATATTTTCAAAATTTAAACACAGATAATTCGTTAA 1609
Db 42174 TTGTATCTAATATAAATTTCTATATATATATATATATATATATATATATATATATATAT 42127

RESULT 7
US-09-949-016-12378/c
; Sequence 12378, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12378
; LENGTH: 119153
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(119153)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12378

Query Match      5.7%; Score 114.2; DB 4; Length 119153;
Best Local Similarity 48.4%; Pred. No. 9.6e-11;
Matches 474; Conservative 0; Mismatches 488; Indels 17; Gaps 5;

QY 530 TCTAAATTTTCATATAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 589
Db 87184 TATATATTTGATCTAGAACAAAGTATCTGGGTGTTTAGCAAAATTTTAAATATACT 87125

QY 590 TTCAAATTAATTTATATGATATGTAAGTACATTAATTTTGTCTATATATCTTTAAATAT 649
Db 87124 TAAATATATATATATATATTTATGAA-TAAACTATATATTTATATAATAATAATATGTA 87066

QY 650 TTCTAAATTAATTAATAAAGCTGATATGAAAAATTTATTTCTTTTAAAGCTATATCAAT 709
Db 87065 TATCATATAATTAATAATATATTAATAATATATAATAATAATAATAATAATAATAATAATA 87006

QY 710 TTATATATATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 768
Db 87005 TTATAATTAATGTAACATATATTAATAATAATAATAATAATAATAATAATAATAATAATAAT 86946
```

```
Qy 769 TTTTGTAAATTTTATTTATCAATTTTAAATAAATTTTACTTTATATGTTTTTTCACAT 828
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86945 ATAAGTATATAATATAAATGTAATTAATATATAATAATTAATTAATTAATGTTATTAAT 86886
Qy 829 TTTTGTAAACAATCATATCATATTAATGTTGAAAGAGAGAAATTTGACAGTGAGTAATAA 888
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86885 ATATAATATATAATTAATTAATCTTATATAATAATATACTTTATATAAAGAGATACATATA 86826
Qy 889 GTGATGAGAAAAAATGTTTATTTCTTAAAAAAAAGCTTAAACAAACATGTAATCACTCT 948
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86825 TATAGTATATATAATATAATAATATATCTTATAATTAATATAATAATATAATATAATAA 86766
Qy 949 CTATTTCACTATCTCTCATTTCAATTTTCTCTTTATCTCTTTTCTTTTATTTTATFAT 1008
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86765 TATATTTTATATATTTTATATATTTTATATATTTTATATATATTTTATATATTTTATAA 86706
Qy 1009 ATCATTTTCAGATTAATTTATTTTACTCTCTTTATTTTCTCTCTATCCCTCTCT----- 1063
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86705 ATAAAAATATATAATATATATATTTTATGTAATTAATTTTATATATACCTTATATAATAT 86646
Qy 1064 -----TATTTCCACTCATATATACACTCCAAAAATTTGGGCATGCTTTTATCACTACTCT 1117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86645 AATATATATATATATTTAATTAATAGTATATAATATATATATATATATATATATATATAC 86586
Qy 1118 ATCTCTCCACTAAATCAATTTAAATGAAGCTGAAAGCATTTGGCAAGCTCTCTCCCTCC 1177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86585 TATGTAATATATATCTATGTAATATATATACATAGTAATAATATGTAATAATGTAATTA 86526
Qy 1178 TCAAGTATTTCCAACTCAGCATTTGGCATCTGATTTGATTTGATATATCTATTTGCAATGTT 1237
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86525 TAATGTAATATATATATATATATATATATATATATATATATATATATATATATATAT-T 86467
Qy 1238 AAAAGTCTTTCCCAATACATAAATTAATTAATCTTAAATAAATAAGGATAAAAATAT 1297
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86466 ATATACTTATATATAGTATCTCTTTATATAATAAGTATATAATATATATATATATATTT 86407
Qy 1298 TTTTCTTTCTTCAAAAAATAAAAATATGTTATTTTGTGTTTGTAGATATATTCGAATAAA 1357
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86406 AATATATATTTTATACATATATTAATATATATATATATATATATATATATATATATAT 86347
Qy 1358 TCTAAATATATGATTAATGATTTTATATATGATTAACATATATCAATATTAATATGA 1417
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86346 ATATTATATATTAAGTATATATATATATATATATATATATATATATATATATATATAT 86287
Qy 1418 TATTTTATATAGTTGTACACATAATTTTAAAGGATAAAAAATATGATAAAAAATAA 1477
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86286 TA---TAATATAAATAATATCAATATATATATATATATATATATATATATATATATAA 86230
Qy 1478 ATTTTAAATATTTTATAT 1496
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86229 GTATATATAAATAATATAT 86211
```

```
RESULT 8
US-09-949-016-16284
; Sequence 16284, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16284
```

```
; LENGTH: 29717
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(29717)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16284

Query Match 5.5%; Score 110; DB 4; Length 29717;
Best Local Similarity 53.3%; Pred. No. 4.3e-10;
Matches 324; Conservative 0; Mismatches 275; Indels 9; Gaps 4;

Qy 493 AATATTGTTTATATCTTAAATTAATAATATTGTCCTTCTAAATTTTCATATAGTTAAAT 552
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15340 ATTTATATATATATTTTATATATATATATATATATATATATATATATATATATATATAT 15399
Qy 553 ATTATATTACTTTTCTCTATCTATCTATTTCTATTTTCAAAATTTATTTATGATAT 612
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15400 ATATTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 15459
Qy 613 GTAAAG-TACATTATATATTTTGCCTATATATCTTAAATTTCTAAATTTTAAAAAAGAC 671
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15460 ATTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 15519
Qy 672 TGATATGAAAAATTTATCTCTTTTAAAGCTATATCAATTTTATATATATATCTTTTCT 731
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15520 ATATATTTATATCTATATTTTATATATATATATATATATATATATATATATATATATAT 15579
Qy 732 TTTCTTTTCAT-TTTCTATTTCAATTTTAAAGAAATAAATTTTCTAAATTTTATTTATCA 790
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15580 ATTTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 15639
Qy 791 ATTTTAAAAATATTTTACTTTTATATATGTTTTCACATTT---TTGTTAAACAAATCAT 846
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15640 ATATATATACATATTTTATATATATATATATATATATATATATATATATATATATATAT 15699
Qy 847 ATCATTTATGATTGAAAGAGAGGAAATTTGACAGTGAGTAATAGTGATGAGAAAAAATGT 906
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15700 ATATTTTATATTTTATATATATATATATATATATATATATATATATATATATATATAT 15759
Qy 907 GTTATTTCTCTAAA---AAAAACCTTAAACAAACATGATCTACTCTCTATTTCTATCT 963
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15760 ATATTTTATATATATTTTATATATATATATATATATATATATATATATATATATATAT 15819
Qy 964 CTCATTTTCATTTTCTCTTATCTCTTTTCTTTTATTTTATCATATCATTTTACATATAA 1023
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15820 ATATTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 15879
Qy 1024 TTATTTTACTCTCTTTTATTTTCTCTCTATCCCTCTCTTTATTTTCCACTCATATACA 1083
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15880 ATATTTATATATATTTTATATATATATATATATATATATATATATATATATATATAT 15939
Qy 1084 CTCCTAAA 1091
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15940 TACCTAAA 15947
```

## RESULT 9

```
US-09-949-016-12423
; Sequence 12423, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12423  
; LENGTH: 60376  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(60376)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12423

Query Match 5.5%; Score 110; DB 4; Length 60376;

Best Local Similarity 53.3%; Pred. No. 4.6e-10;

Matches 324; Conservative 0; Mismatches 275; Indels 9; Gaps 4;

Qy	493	AATATGTTGTTATATCTTAAATTAAGAAATATGTCCTCTCTAAATTTTCATATAGTTAAT	552
Db	47379	ATTATATATATATTTTATATATTTATATATATATTTATATATTTATATATATTTATAT	47438
Qy	553	ATTATATATCTTTTCTCTATCTTCTATTTCTATTTCTAAATTTATTTATGTCATAT	612
Db	47439	ATATTTATATTTATATATATTTATATTTATTTATTTATATATTTATTTATATATAT	47498
Qy	613	GTAAG-TACATATATATTTTGTCTATATACATTAATATTTCTAAATTTAAAAAGAC	671
Db	47499	ATTATATATATTTATATATATATTTATATATATTTATATATATTTATATATATTTAT	47558
Qy	672	TGATATGAAGAAATTTATCTTTTAAAGCTATATCAATTTTATATATACITTTCTTTCT	731
Db	47559	ATATATTTATATGATATATTTTATATATTTTATATATTTTATATATATTTATATAT	47618
Qy	732	TTTCTTTTCAT-TTCTTATTCATTTTAAAGAAATTAATTTTGTAAATTTTATTTATCA	790
Db	47619	ATTTTATATTTATATATACATATTTCAATATATTTTATATATTTTATATATTTAT	47678
Qy	791	ATTTTAAAAATATTTTCTTTATATGTTTCTTCTCAATTT---TTGTTAAACAAATCAT	846
Db	47679	ATATATATACATATTTTATATATATATATATATTTATATATATACATATTTTATAT	47738
Qy	847	ATCATTTATGTTGAAGAGAGAAATTCACAGTGTAGTAATAGTATGATGAGAAAAATGT	906
Db	47739	ATATTTATATTTATATATATTTTATATATATTTTATATATTTTATATATTTTATAT	47798
Qy	907	GTTATTTCTCTAAA---AAAAACCTAAACACATGATCTCTCTATTTTCATCTATCT	963
Db	47799	ATATTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATAT	47858
Qy	964	CTCATTTCTATTTCTCTTTATCTCTTTTATTTTATTTTATATATATTTTATATAT	1023
Db	47859	ATATTTATATTTATATATTTTATATATTTTATATATTTTATATATTTTATATAT	47918
Qy	1024	TTATTTTACTCTCTTTATTTTCTCTATCTATCTCTCTTTATTTTCCACTCATATATACA	1083
Db	47919	ATATTTATATATTTTATATTTTATATATATTTTATATATTTTATATATTTTATAT	47978
Qy	1084	CTCCAAA 1091	
Db	47979	TACCTAAA 47986	

RESULT 10

US-09-949-016-17145  
; Sequence 17145, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17145  
; LENGTH: 30820  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17145

Query Match 5.4%; Score 109.4; DB 4; Length 30820;

Best Local Similarity 47.0%; Pred. No. 5.4e-10;

Matches 470; Conservative 0; Mismatches 521; Indels 8; Gaps 4;

Qy	487	AAAAGAAATATTTGTTATATTTCTTAATTAAGAAATATGTCCTCTCTAAATTTTCATATAG	546
Db	5863	AAAATATATATATATAAATATATATATTAATTAATTAATTAATTAATTAATTAATATAT	5922
Qy	547	TTAATTTATATATCTTTTCTCTATCTTCTATTTCTAGTTCTATTTTCAAATTTATTTAT	606
Db	5923	ATTATATATATATTTATATATATATTAATTAATTAATTAATTAATTAATTAATTAATAT	5982
Qy	607	GCATATGTAAAGTACATTTATATTTTGTCTATATATCTTTAAATATTTTCTAAATTTTAAAA	666
Db	5983	AAATATATATATATAAATATATATATTAATTAATTAATTAATTAATTAATTAATTAATATA	6042
Qy	667	AAGCTGATATGAAAAATTTATCTTTTAAAGCTATATCAATTTTATATATACITTTTCT	726
Db	6043	TATAT---TATATACTATATATAAATAATATAATTAATTAATTAATTAATTAATTAATAT	6099
Qy	727	TTTCTTTCTTTCTATTTCTCAATTTCAATTTAATAAGAAATTAATTTTGTAAATTTTATTT	786
Db	6100	AAATATATATTTATATATCTATATATAATTAATTAATTAATTAATTAATTAATTAATTAAT	6159
Qy	787	ATCAATTTATA-AAAAATATTTTACTTTTATATGTTTCTTTTTCACATTTTCTTTAAACAAATCA	845
Db	6160	ATAATATATATTTATATATCTATATATATAATTAATTAATTAATTAATTAATTAATTAAT	6219
Qy	846	TATCATTTATGTTGAAGAGAGAAATTTGACAGTGTAGTAATAGTATGATGAGAAAAATG	905
Db	6220	TTATATCTATTTATATAGTATAATATAGAATATATTTATATATTTTATATATAAGATATA	6279
Qy	906	TGTTATTTCTTAAAAAAACCTTAAACAAACATGATCTCTCTCTCTATTTTCAT-CTATCTC	964
Db	6280	TATTTATATTTGTATATAAGATATAATATATTTATATTTTGTATATAAGATATAATATAT	6339
Qy	965	TCATTTTCATTTTCTCTTTTATCTCTTTTATTTTATTTTATCATATCATTTTCACATTAAT	1024
Db	6340	ATATATTTGTATATAAGATATAATATATTTATTTATTTATTTATTTATTTATTTATATAT	6399
Qy	1025	TATTTTCTCTCTTTATTTTCTCTATCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1084
Db	6400	TGTTATATAAGATATAATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	6459
Qy	1085	TCCAAATTTGGGCGATCGCTTTTATCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1144
Db	6460	TATAAGATATAATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	6519
Qy	1145	AACTGAAAAGCATTTGGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1204
Db	6520	AGATATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	6579
Qy	1205	ATCTGATTTGATTCAGTAT---ATCTATTTGCGATGTTAAAGTCTTTTCCCAATACATAAC	1261
Db	6580	ATATTTATATTTGATATAAGATATAATATATTTATTTATTTATTTATTTATTTATTTAT	6639
Qy	1262	TATTAATTAATCTTAAATAAATAAGGATAAATAATTTTCTTTCTTCTTCAATAAATAAATA	1321



```
QY 994 TTATTTTTCATCATCATCTTTTCAATTAATTTTCTACTCTCTTTTCTCTCTCT 1053
Db TACAATATATATATATATATCTTATATACAAATATATATATATATATCTTATATACA 6428
QY 1054 ATCCCTCTCTTATTTCCACTCATATACACTCCAAATTTGGGCGATGCGCTTTATCACTA 1113
Db ATATATAATATATATATATCTTATATACAAATATATATATATATATATATATATATATAT 6379
QY 1114 CTCTATCTCTCCATCAATCAATTTAAATGAAACCTGAAAGCATTTGGCAAGTCTCTCC 1173
Db ATATACAAATATATATATATATCTTATATACAAATATATATATATATATATATATATATAT 6319
QY 1174 CTCCTCAAGTGATTTCCAACTCAGCATTTGGCATCTGATTTAGTATATCAATCTATTTGCAT 1233
Db AATATATATATATATATATCTTTATATACAAATATATATATATATATATATATATATATCT 6262
QY 1234 GTGTAAGAGTCCTTTCCCAATACATAAATCTTAAATTTAACTTTAAATAAATAAGGATAAA 1293
Db ATATATATATATCTATATATATATATATATATATATATATATATATATATATATATATAT 6202
QY 1294 ATATTTTCTTTCTTATATATATATATATATATATATATATATATATATATATATATAT 1353
Db GTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6142
QY 1354 TAAATCTAAATATATATATATATATATATATATATATATATATATATATATATATAT 1413
Db AATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6082
QY 1414 ATGATATATTTTATATATATATATATATATATATATATATATATATATATATATATAT 1473
Db ATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6022
QY 1474 ATAAATTTTAAATTTTATATATATATATATATATATATATATATATATATATATATAT 1533
Db ATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5962
QY 1534 ACCGATATTTTACAACTTATATATATATATATATATATATATATATATATATATATAT 1593
Db ATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5902
QY 1594 AACGATATAT 1603
Db TATATATATAT 5892
```

## RESULT 13

```
US-09-949-016-15348
; Sequence 15348, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15348
```

```
; LENGTH: 134987
; TYPE: DNA
```

```
; ORGANISM: Human
US-09-949-016-15348
```

## Query Match

```
Best Local Similarity 50.4%; Score 105.4; DB 4; Length 134987;
Matches 284; Conservative 0; Mismatches 276; Indels 3; Gaps 1;
```

```
QY 487 AAAAGAAATATGTTTATATATCTTAAATTTAAATTTTCCCTTCTAAATTTTTCATATAG 546
Db ATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11458
QY 547 TTAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 606
Db ATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11518
QY 607 GCATATGTAAAGTACATTTATATATATATATATATATATATATATATATATATATATAT 666
Db TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11518
QY 667 AAGACTGATATGAAAAATTTTATCTTTTAAAGCTATATCAT---TTTATATATATATAT 723
Db ATTTATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 11638
QY 724 TCTTTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 783
Db TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11698
QY 784 TTTATCAATTTTAAATATATATATATATATATATATATATATATATATATATATATATAT 843
Db TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11758
QY 844 CATATCATTTATGATTTGAAAGAGAGGAAATTCACAGTGTAGTAATAAGTCATGAGAAAAA 903
Db TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 11818
QY 904 TGTCTTTATTTCTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 963
Db TGTATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 11878
QY 964 CTCAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1023
Db TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 11938
QY 1024 TTATTTTCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1046
Db TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 11939
```

## RESULT 14

```
US-09-949-016-15349
```

```
; Sequence 15349, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15349
```

```
; LENGTH: 134987
; TYPE: DNA
```

```
; ORGANISM: Human
US-09-949-016-15349
```

## Query Match

```
Best Local Similarity 50.4%; Score 105.4; DB 4; Length 134987;
Matches 284; Conservative 0; Mismatches 276; Indels 3; Gaps 1;
QY 487 AAAAGAAATATGTTTATATATCTTAAATTTAAATTTTCCCTTCTAAATTTTTCATATAG 546
```

```
Db 11399 ATATTATATATATATATATATATATATATATATATATATATATATATATATATATAT 11458
Qy 547 TTAATTATATATATATATATATATATATATATATATATATATATATATATATATATAT 606
Db 11459 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11518
Qy 607 GCATATGTAAGTACATATATATATATATATATATATATATATATATATATATATATAT 666
Db 11519 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11578
Qy 667 AAGCTGATATGAAAAATTTTCTTCTATATATATATATATATATATATATATATATATAT 723
Db 11579 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11638
Qy 724 TCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 783
Db 11639 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11698
Qy 784 TTTATCAATTTTAAAAATATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 843
Db 11699 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11758
Qy 844 CATATCAATATGATGTAAGTACATATATATATATATATATATATATATATATATATATAT 903
Db 11759 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11818
Qy 904 TGCTGTTATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 963
Db 11819 TGTATAATATATATATATATATATATATATATATATATATATATATATATATATATAT 11878
Qy 964 CTCATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1023
Db 11879 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11938
Qy 1024 TTAATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1046
Db 11939 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11961
```

RESULT 15  
US-09-949-016-15350  
; Sequence 15350, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15350  
; LENGTH: 134987  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15350

Query Match 5.2%; Score 105.4; DB 4; Length 134987;  
Best Local Similarity 50.4%; Pred. No. 3.2e-09;  
Matches 284; Conservative 0; Mismatches 276; Indels 3; Gaps 1;  
Qy 487 AAAAGAATATGTTTAT 546  
Db 11399 ATATTAT 11458  
Qy 547 TTAATTAT 606

Search completed: September 1, 2005, 22:18:37  
Job time : 254.78 secs

```
Db 11459 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11518
Qy 607 GCATATGTAAGTACATATATATATATATATATATATATATATATATATATATATATAT 666
Db 11519 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11578
Qy 667 AAGACTGATATGAAAAATTTTCTTTTAAAGCTATATCAT---TTTATATATATACCTTTT 723
Db 11579 ATTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11638
Qy 724 TCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 783
Db 11639 TATATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 11698
Qy 784 TTTATCAATTTTAAAAATATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 843
Db 11699 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11758
Qy 844 CATATCAATATGATGTAAGTACATATATATATATATATATATATATATATATATATATAT 903
Db 11759 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11818
Qy 904 TGCTGTTATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 963
Db 11819 TGTATAATATATATATATATATATATATATATATATATATATATATATATATATATAT 11878
Qy 964 CTCATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1023
Db 11879 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11938
Qy 1024 TTAATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1046
Db 11939 TATATTTTATATATATATATATATATATATATATATATATATATATATATATATAT 11961
```